# Quasi-Periodic Stability of Subfamilies of an unfolded skew Hopf bifurcation

Broer, HW; Wagener, FOO

*Published in:*
Archive for Rational Mechanics and Analysis



*Publication date:*
2000

Arch. Rational Mech. Anal. 152 (2000) 283–326. © Springer-Verlag 2000

# *Quasi-Periodic Stability of Subfamilies of an Unfolded Skew Hopf Bifurcation*

H. W. BROER & F. O. O. WAGENER

*Communicated by* M. GOLUBITZKY

## Abstract

In the skew Hopf bifurcation a quasi-periodic attractor with nontrivial normal linear dynamics loses hyperbolicity. The simplest setting concerns rotationally symmetric diffeomorphisms of $S^1 \times \mathbf{R}^2$. Their dynamics involve periodicity, quasi-periodicity and chaos, including mixed spectrum. The present paper deals with the persistence under symmetry-breaking of quasi-periodic invariant circles in this bifurcation. It turns out that, when adding sufficiently many unfolding parameters, the invariant circle persists for a large Hausdorff measure subset of a submanifold in parameter space.

## Contents

## 1. Introduction

We begin this introduction with a quick sketch of the mathematical result of this paper, which is a perturbation theorem of Kolmogrov-Arnol'd-Moser-type. Then the main application is introduced, after which the contents of this paper are explained at a more leisurely pace and in more detail.

**Main result.** This paper considers parametrized families of diffeomorphisms $\varphi_p$ of $M = S^1 \times \mathbf{R}^2$ to itself, of the following form:

$$\varphi_p(x, y) = (x + \omega(p) + \varepsilon f(x, y, p), \beta(p) E_k(x) y + \varepsilon g(x, y, p)) , \qquad (1)$$

where the phase variables $x$ and $y$ are in $S^1 = \mathbf{R}/2\pi\mathbf{Z}$ and $\mathbf{R}^2$ respectively, the parameter $p$ takes values in an open subset $P$ of $\mathbf{R}^q$ (called the space of parameters of $\varphi$), $\omega(p)$ and $\beta(p)$ depend in a real analytic way on $p$, and the perturbation strength $\varepsilon$ is some real number $0 \leqq \varepsilon \leqq 1$. The functions $f(x, y, p) \in \mathbf{R}$ and $g(x, y, p) \in \mathbf{R}^2$ are assumed to be real analytic, and $E_k(x) \in \mathrm{GL}(2, \mathbf{R})$ equals

$$E_k(x) = \begin{pmatrix} \cos kx & -\sin kx \\ \sin kx & \cos kx \end{pmatrix},$$

where $k \in \mathbf{Z}\backslash\{0\}$.

For $\varepsilon = 0$, the circle $S = S^1 \times \{0\}$ is invariant under $\varphi$. The main result of this paper implies that under certain non-degeneracy conditions on $\omega(p)$ and some regularity assumptions on $f$ and $g$, there is a constant $\varepsilon_0$ such that persistence (in a sense to be specified) of an invariant circle for $0 \leqq \varepsilon \leqq \varepsilon_0$ is at most a codimension $6|k|$ phenomenon. See for instance Arnol'd [1] for this terminology.

In particular, it will be shown that there exists an unfolding $\tilde{\varphi}_{p,\sigma}$ of $\varphi_p$ (that is, a larger family that satisfies $\tilde{\varphi}_{p,0} = \varphi_p$) such that in the parameter space $\tilde{P}$ of the unfolding there exists a smooth manifold of codimension $6|k|$, and a subset of positive measure in that submanifold, such that for parameters in that subset the corresponding diffeomorphism has an invariant circle homotopic to $S^1 \times \{0\}$.

**Background.** This result should be compared with the well-known theory of small perturbations of systems of the following form:

$$\psi_p(x, y) = (x + \omega(p) + \varepsilon f(x, y, p), A(p)y + \varepsilon g(x, y, p)),$$

where $x$, $y$ and $p$ are as above, and where $\omega(p) \in \mathbf{R}$ and $A(p) \in \mathrm{GL}(2, \mathbf{R})$ depend in a real analytic way on $p$. Here the invariant tori $S^1 \times \{0\}$ of the unperturbed ($\varepsilon = 0$) system have an $x$-independent normal linear part. Systems with this property are said to be of Floquet form, and systems which can be brought into this form by a coordinate transformation are said to be reducible (to Floquet form). Note that the set-up is chosen analogously to (1). Here the result is (compare [4]) that under nondegeneracy conditions on $\omega(p)$ and $A(p)$ there is an unfolding $\tilde{\psi}_{p,\sigma}$ of $\psi_p$, and a set of positive measure in the parameter space of the unfolding, such that for parameters in that set the system has an invariant circle.

A well-known argument given below, shows that systems of the form (1) are not reducible to Floquet form, due to topological obstructions. To our knowledge, it has not been previously known whether persistence of invariant circles for systems of the type (1) has finite codimension, and as far as we know this paper is the first successful attempt to develop KAM theory for such a non-reducible case. Technically the difference with the "classical" theory is that the linearized conjugacy equations are coupled instead of decoupled. This difficulty is overcome by the introduction of many extra parameters. After this, our approach turns out to be an adaptation of the "classical" case, see for instance [3], [4].

**Main application.** The main motivation to study the system (1) is the application of the results to the *skew Hopf bifurcation* of an invariant quasi-periodic circle in a family of diffeomorphisms on $S^1 \times \mathbf{R}^2$. It is a variation on the "quasi-periodic" or "reducible" Hopf bifurcation (see [2], [3], [4]) of a quasi-periodic circle attractor to a quasi-periodic torus attractor. In the skew Hopf bifurcation, a quasi-periodic circle attractor bifurcates to a "weakly chaotic" attractor.

The class of skew Hopf bifurcations can be introduced starting out from *integrable* bifurcations: these are special bifurcations, possessing a rotational symmetry, which makes them relatively easy to analyse. We shall consider small non-symmetric perturbations of such an "integrable" skew Hopf family and study their effects.

The integrable skew Hopf family appeared in [6], where a bifurcation analysis of rotationally symmetric skew Hopf families was given. The analysis in [6] consists of the following two steps:

1. Using symmetry considerations, the existence of an invariant circle is shown. Dynamics on the circle are resonant or quasi-periodic.
2. The existence of an invariant 2-torus, bifurcating from the invariant circle, is shown by analysing the dynamics in a neighbourhood of the invariant circle.

The genesis of an invariant 2-torus motivates the name skew *Hopf* bifurcation.

**Problem.** In this paper, arbitrary, non-symmetric, small perturbations are applied to an integrable skew Hopf family. The first step of the above program leads to the following question:

> *Under what conditions does an invariant quasi-periodic circle persist in a skew Hopf bifurcation family?*

It turns out that an *unfolding* of the skew Hopf family has to be considered, and that the invariant circle persists for subfamilies of that unfolding. To obtain the result, an extension of ordinary KAM theory is developed, which is the main mathematical result of this paper. The remainder of the program, the matter of invariant 2-tori, is considered elsewhere (see [15]; for an overview, also see [7]).

**History.** The skew Hopf bifurcation was introduced by CHENCINER & IOOSS [8], [9], as an alternative to the *reducible* (quasi-periodic) Hopf bifurcation (see [2], [3], [4]). BROER & TAKENS (in [6]) were interested in the phenomenon that, in the (rotationally symmetric) case they considered, an invariant torus persists, carrying ergodic dynamics having a mixed (i.e., pure point *and* continuous) spectrum. This paper consideres the effects of small, non-symmetric perturbations of the symmetric skew Hopf system.

### *1.1. Integrable system*

The natural phase space of a skew Hopf family (after centre manifold reductions, see [10], [12]) is $S^1 \times \mathbf{R}^2$, where $S^1 = \mathbf{R}/2\pi\mathbf{Z}$. The general form of the integrable skew Hopf family is (compare [6]):

$$(x, y) \mapsto \Big(x + \omega + |y|^2 f(|y|^2, p), \Big(\beta + |y|^2 g(|y|^2, p)\Big) E_k(x)y\Big), \tag{2}$$

where the phase variables $x$ and $y$ take values in $S^1$ and $\mathbf{R}^2$ respectively; the parameters $\omega$, $\beta \in \mathbf{R}$, $p \in P \subset \mathbf{R}^q$; the norm $|.|$ is defined as $|y|^2 = y_1^2 + y_2^2$; the functions $f$ and $g$ take values in $\mathbf{R}$; finally, $E_k(x)$ is defined for $k \in \mathbf{Z}\backslash\{0\}$ as

$$E_k(x) = \begin{pmatrix} \cos kx & -\sin kx \\ \sin kx & \cos kx \end{pmatrix}.$$

### *1.2. Non-reducibility*

The integrable skew Hopf bifurcation family cannot be isotopic to a diffeomorphism of Floquet form in the class of diffeomorphisms. This can be seen most readily by introducing the concept of the *linking number* of two circles. First this is introduced for disjoint circles $S_1$, $S_2$ of the following special form:

$$S_j = \Big\{(x, g_j(x)) \in S^1 \times \mathbf{R}^2\Big\}, \tag{3}$$

where $g_j : S^1 \to \mathbf{R}^2$ are continuous functions. Disjointness of the circles as point sets is equivalent to the requirement that $g_1(x) \neq g_2(x)$ for all $x \in S^1$. Define a function $f(x)$ by:

$$f(x) = \frac{g_2(x) - g_1(x)}{|g_2(x) - g_1(x)|}.$$

The map $f(x)$ takes values in $\{x \in \mathbf{R}^2 : |x| = 1\}$, which is diffeomorphic to $S^1$; hence $f$ maps $S^1$ to itself. The linking number $\ell(S_1, S_2)$ of $S_1$ and $S_2$ is defined to be the degree of $f$; intuitively speaking, this is the number of times $f(x)$ performs a complete revolution around 0, taking into account the orientation.

For instance, if $S_1$ and $S_2$ are given by:

$$g_1(x) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{and} \quad g_2(x) = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

then it is readily verified that $\ell(S_1, S_2) = 0$. Now, consider the linking number of $\varphi(S_1)$ and $\varphi(S_2)$, the images of $S_1$ and $S_2$ under application of the integrable skew Hopf map (2). Note that $S_1$ is invariant: $\varphi(S_1) = S_1$. Also note that $\varphi(S_2)$ is given by $\tilde{g}_2$, where

$$\tilde{g}_2(x) = \begin{pmatrix} \cos k(x - \omega - f(1, p)) \\ \sin k(x - \omega - f(1, p)) \end{pmatrix}.$$

It is readily seen that $\ell(\varphi(S_1), \varphi(S_2)) = k$. On the other hand, if $\psi$ is a linear diffeomorphism of Floquet form, that is,

$$\psi(x, y) = (x + \omega, Ay),$$

it follows that $\ell(\psi(S_1), \psi(S_2)) = 0$.

Without proof, we mention the topological fact that the linking number of circles homotopic to those of the form (3) is invariant under isotopy in the class of diffeomorphisms. If the integrable skew Hopf system could be brought into Floquet form by a coordinate transformation, the linking number of $\varphi(S_1)$ and $\varphi(S_2)$ should be zero. But it has been shown that it is equal to $k \neq 0$, and the statement at the beginning of this subsection follows.

Actually, the topological fact also implies that small perturbations of an integrable skew Hopf family cannot be isotopic to a Floquet diffeomorphism in the class of diffeomorphisms. Hence, these diffeomorphisms cannot be reduced to Floquet form as well.

### *1.3. Appearance of unfolding parameters*

If small, non-symmetric perturbations are applied to an integrable skew Hopf family, it may be asked whether they can be “killed” by well-chosen coordinate transformations. Postulating equality between a transformed system (where the transformation is as yet unknown) and a “nice” system (integrable, or at least integrable if truncated to lowest order terms), gives rise to a complicated nonlinear conjugacy equation. The general approach of KAM theory is to solve this equation by succesively linearizing it at some starting guess for the unknowns, using the solutions of the linear equation as a better guess for the unknowns, and linearizing again *ad infinitum*.

In the context of the skew Hopf bifurcation family, the most important linearized homological equation has the following structure:

$$v(x + \omega) - e^{ikx} v(x) = g(x). \tag{4}$$

Here $v(x) \in \mathbf{C}$ corresponds to the coordinate transform, $g(x)$ to the perturbation to be transformed away, and $e^{ikx}$ corresponds to the matrix $E_k(x)$; $k \in \mathbf{Z}\backslash\{0\}$. In [6] this equation was shown to have a $|k|$-dimensional complex obstruction (equivalent to a $2|k|$-dimensional real one).

Intuitively, this means the following. Let $\sum_{n\in\mathbf{Z}} g_n e^{inx}$ be the Fourier expansion of $g(x)$, and consider the coefficients $g_n$, $n \notin \{1, \ldots, k\}$ as given. Then there are unique values $g_1^*, g_2^*, \ldots, g_k^*$ which the coefficients $g_1, g_2, \ldots, g_k$ have to take in order that (4) has a real analytic (or even continuous) solution $v(x)$.

Since generically the coefficients $g_1, \ldots, g_k$ will differ from these special values, in the general case a small, non-symmetric perturbation cannot be transformed away.

However, if instead of $g(x)$ a function $G(x, \sigma)$ is put in the right-hand side, such that $G(x, \sigma)$ equals $g(x)$ save for the fact that the Fourier coefficients $G_1(\sigma)$, $\ldots$, $G_k(\sigma)$ are of the form

$$G_1(\sigma) = g_1 + \sigma_1, \ldots, G_k(\sigma) = g_k + \sigma_k,$$

then (4) has a solution for $G(x, \sigma^*)$, where

$$\sigma_1^* = g_1^* - g_1, \ldots, \sigma_k^* = g_k^* - g_k.$$

The family $G(x, \sigma)$ is called an unfolding of $g(x)$, since $g(x)$ is a (zero-dimensional) subfamily of $G(x, \sigma)$:

$$g(x) = G(x, 0).$$

Put in another way, if a $|k|$-dimensional complex ($2|k|$-dimensional real) unfolding parameter $\sigma$ is added (in the right way), there is one value of the parameter for which (4) has a solution. Roughly, the proof of the main theorem works because unfolding parameters have to be added only once, and not at every step.

Note that if the original right-hand side already depends on a $q$-dimensional real parameter $p$ such that the map

$$p \mapsto (g_1(p), \ldots, g_k(p)),$$

has an injective derivative, then only $2|k| - q$ real parameters have to be added to obtain a solution of (4).

The main result of this paper can now be described as follows. If a suitable unfolding of the original perturbed skew Hopf system is considered, there is a submanifold of codimension $6|k|$ in the space of parameters (denoted by $\mathcal{N}$) and, on that submanifold, there is a subset of large measure (denoted by $\mathcal{N}_c$), which is diffeomorphic to the Cartesian product of a Cantor set and a manifold (called Cantor-structured for short): see Fig. 1. For parameters in this Cantor-structured set, the invariance of a quasi-periodic (actually Diophantine) circle can be shown. The Cantor-structure reflects the fact that a Diophantine condition has to be imposed on the rotation number of the circle dynamics.

Not all of the non-symmetric perturbation is transformed away, only the parts independent of and linear in $y$. Note however that this is sufficient to show the persistence of an invariant circle.

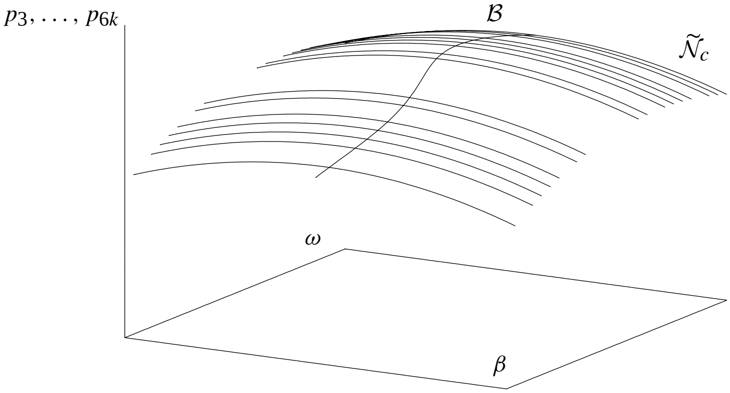


**Fig. 1.** A sketch of the parameter space $P$. Indicated is the large measure subset $\tilde{\mathcal{N}}_c$ of the codimension $6|k|$-submanifold $\tilde{\mathcal{N}}$ in $P$. For $p \in \tilde{\mathcal{N}}_c$ the family of diffeomorphisms $\varphi_p$ has an invariant quasi-periodic circle. Also indicated is the "bifurcation curve" $\mathcal{B}$, where the invariant circles are not normally hyperbolic. Note that this picture also applies to for the non-skew case, where $k = 0$.

Attention is drawn to the allowed perturbations: these are required to be real analytic with a complex analytic extension to a complex strip of width $r_0$ around the real domain. The size of the allowed perturbation is then $\mathrm{O}(r_0)$. We do not know whether this is a typical requirement for this type of problems, or whether it can be eventually removed. Without elaborating, we point out the fact that this prevents us extending the present result to the case of finite differentiability.

### *1.4. Fattening*

It is well known from the theory of invariant manifolds (see [12]), that normal hyperbolicity is an open property: in particular, if a parametrized family of dynamical systems has a normally hyperbolic invariant manifold at a parameter $p^*$, then in the space of parameters $P$ there is an open neighbourhood of $p^*$ such that the system has normally hyperbolic invariant manifolds for all parameters in that neighbourhood.

The result which is obtained in this paper allows more to be said about the structure of the open subset $\mathcal{F}$ of $P$ for which the existence of (normally hyperbolic) invariant circles can be shown. First, attention is restricted to the codimension $6|k|$ submanifold mentioned in the previous subsection. One of the consequences of the main result is that the change of coordinates that transforms away the zeroth and first order terms of the original (non-symmetric) perturbations for parameters in the Cantor-structured subset, is in fact infinitely differentiable in the parameters. After transformation, the system is still of the general form of an (non-symmetrically) perturbed integrable skew Hopf system; where, however, the two lowest order perturbation terms vanish on the Cantor-structured subset. Since a Cantor set is a perfect set, every point in it is the accumulation point of other points in the set. From this it follows that, together with the perturbations, all their derivatives vanish on the Cantor-structured subset as well: the perturbations are there flat in $p$.

The regions of normal attraction and normal repulsion touch at a bifurcation point $p_0$. The main result yields that the set $\mathcal{B}$ of these bifurcation points is also Cantor-structured, but of one dimension less than $\mathcal{N}_c$. The flatness of the (transformed) perturbation implies that the order of contact of the two regions of normal hyperbolicity, restricted to the codimension $6|k|$ submanifold $\mathcal{N}$, is infinite. These two regions can be considered as the intersection of $\mathcal{N}$ with $\mathcal{F}$.

Normal to $\mathcal{N}$, the set $\mathcal{F}$ might be in general rather thin. In [15], a model system is considered to get an impression of the shape of $\mathcal{F}$. This system is of the form:

$$\varphi_p(x, y) = \left(x + \omega, \left(\beta - |y|^2\right) E_k(x)y + \begin{pmatrix} \varepsilon & 0 \\ 0 & -\varepsilon \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} \varepsilon \\ 0 \end{pmatrix}\right).$$

Here $x \in S^1$, $y \in \mathbf{R}^2$, $\omega \in \mathbf{R}$, $\beta > 0$ and $0 \leqq \varepsilon < 1$. It is shown that for $|\varepsilon| < 1 - \beta$, the map $\varphi_p$ has a unique attracting invariant circle. For $|\varepsilon| < f^{-1}(\beta)$, where $f(x) = 1 + 3\,(x/2)^{2/3} + x$ (thus $f^{-1}(\beta) = c(\beta - 1)^{3/2} + \mathrm{o}((\beta - 1)^{3/2})$), $\varphi_p$ has a unique repelling invariant circle.

## 2. Main theorem

This section presents the main theorem of the present paper. Many of the concepts and ideas in this section are freely quoted from [3], [4].

### *2.1. Definition*

**Phase space.** Let $X$ denote the three-dimensional pahse space $S^1 \times \mathbb{R}^2$, where $S^1 = \mathbb{R}/2\pi Z$, and let the parameter space $P$ be an open subset of a finite dimensional vector space. Typical points in $S^1$, $\mathbb{R}^2$ and $P$ will be denoted by $x$, $y$ and $p$ respectively.

**Diffeomorphisms.** Consider $P$-parametrized real analytic (for a definition of real analyticity see below) families $\varphi_p(x, y)$ of diffeomorphisms on $X$,

$$\varphi_p : X \to X,$$

which are also denoted by $\varphi : X \times P \to X \times P$, with

$$\varphi : (x, y, p) \mapsto \left(\varphi_p(x, y),\ p\right)$$

on $X \times P$. This kind of diffeomorphisms, preserving the $p$-coordinate, will be called *vertical*. Let $\varphi$ be such that near $y = 0$ it has the form

$$\varphi(x, y, p) = (x + a(p) + \chi(x, y, p),\ b(p)E_k(x)y + \psi(x, y, p),\ p)\,, \qquad (5)$$

where $a(p)$, $b(p)$, $\chi(x, y, p) \in \mathbf{R}$ and $\psi(x, y, p) \in \mathbf{R}^2$, and $E_k(x)$ is as in the introduction:

$$E_k(x) = \begin{pmatrix} \cos kx & -\sin kx \\ \sin kx & \cos kx \end{pmatrix},$$

for $k \in \mathbf{Z}\backslash\{0\}$. Here the functions $\chi$ and $\psi$ are of order $O(|y|)$ and $O(|y|^2)$ respectively. Note that consequently the set $S = \{(x, y) \mid y = 0\}$ is an invariant circle. As in the introduction, the factor $b(p)$ "controls" the normal hyperbolicity of $S$. Typically the case where $b(p) \approx 1$ is considered, where the invariant circle fails to be normally hyperbolic. The term $a(p)$ "controls" the rotation number of the diffeomorphism restricted to the circle.

**Essential non-reducibility.** In the introduction, we remarked (Section 1.2) that $\varphi(x, y, p)$ is not isotopic to the identity, that is, it cannot be deformed to the identity map within the class of all diffeomorphisms. This implies that the $x$-dependence of the linear part of the diffeomorphism is essential, and that this Poincaré map cannot be taken from a vector field defined on a Cartesian product $\mathbf{T}^2 \times \mathbf{R}^2$.

**Unfoldings.** In order to formulate our stability result, the following unfolding of the diffeomorphism is considered (also denoted by $\varphi$):

$$\begin{aligned}\varphi(x, y, p) = \Big(&x + a(p) + \chi(x, y, p),\\ &E(x, b(p))y + M(x, m(p)) + L(x, \ell(p))y + \psi(x, y, p),\ p\Big),\end{aligned}$$

where $a(p)$, $\chi(x, y, p)$ and $\psi(x, y, p)$ are as above; but now

$$b(p) = (b_1(p), b_2(p)) \in \mathbf{R}^2,$$

$m(p) \in \mathbf{R}^{2k}$ and $\ell(p) \in \mathbf{R}^{4k}$. The maps $E$, $M$ and $L$ are defined as follows:

$$E(x, b) = \begin{pmatrix} \cos kx & -\sin kx \\ \sin kx & \cos kx \end{pmatrix} \begin{pmatrix} b_1 & -b_2 \\ b_2 & b_1 \end{pmatrix}, \tag{6}$$

$$M(x, m) = \sum_{n=0}^{k-1} \begin{pmatrix} \cos nx & -\sin nx \\ \sin nx & \cos nx \end{pmatrix} \begin{pmatrix} m_{n,1} \\ m_{n,2} \end{pmatrix}, \tag{7}$$

and

$$L(x, \ell) = \sum_{n=-k}^{k-1} \begin{pmatrix} \cos nx & -\sin nx \\ \sin nx & \cos nx \end{pmatrix} \begin{pmatrix} \ell_{k+n,1} & -\ell_{k+n,2} \\ \ell_{k+n,2} & \ell_{k+n,1} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{8}$$

Note that all matrices save one in the above expressions are conformal orientation preserving linear maps; the one exeption is a reflection. For example the matrix $B = \begin{pmatrix} b_1 & -b_2 \\ b_2 & b_1 \end{pmatrix}$ is a conformal linear map in the plane: it can be written as

$$B = \|b\| \begin{pmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{pmatrix},$$

where $\|b\|^2 = b_1^2 + b_2^2$ and $\cos\vartheta = b_1/\|b\|$, $\sin\vartheta = b_2/\|b\|$. By conjugating the diffeomorphism to a map $x \mapsto x + \vartheta(p)/k$, $B$ can be made equal to $\|b\| I$, where $I$ is the identity matrix. Note also that for $b_2(p) = 0$, $m(p) = 0$ and $\ell(p) = 0$ the original diffeomorphism is recovered; so $\varphi$ is indeed an unfolding.

The $M(x, m(p))$ and $L(x, \ell(p))$ are added to deal with *obstructions* arising from the special (twisting) form of the normal behaviour of the diffeomorphism (see [6]). This ultimately leads to results valid on manifolds of positive co-dimension in the parameter space.

**Nondegeneracy.** Let $p_0 \in P$ be given. We say that $\varphi$ is *nondegenerate at* $p_0$, if the map $(a, b, m, \ell) : P \to \mathbf{R} \times \mathbf{R}^2 \times \mathbf{R}^{2k} \times \mathbf{R}^{4k}$ has a surjective derivative at $p_0$.

**Diophantine frequencies.** Define the set $\mathbf{R}_c$ of *Diophantine frequencies* ($c$ as in $C$antor). For given constants $\tau > 2$ and $\gamma > 0$ (which are fixed from now on throughout the whole paper):

$$\mathbf{R}_c = \left\{ \omega \in \mathbf{R} \;\middle|\; \forall p \in \mathbf{Z},\ \forall q \in \mathbf{Z}\backslash\{0\} : \left| \omega - \frac{p}{q} \right| \geqq \gamma |q|^{-\tau} \right\}. \tag{9}$$

If $\Sigma \subset P$ is any open subset, write $\Sigma_c = \Sigma \cap a^{-1}(\mathbf{R}_c)$. If the restriction of the map $a$ to $\Sigma$ is a submersion, then $\Sigma_c$ will be called a *Whitney-smooth family of manifolds, parametrized over the Cantor set* $\mathbf{R}_c$. Such a family will also be called a "Cantor foliation". The rest are subsets $a^{-1}(\omega)$, $\omega \in \mathbf{R}_c$, of the parameter space $P$. Note that if $\gamma > 0$ is sufficiently small, then the set $\Sigma_c \subset P$ has positive measure. By the smoothness of the map $(a, b, m, \ell)$, the diffeomorphim $\varphi(x, y, p)$ is nondegenerate at a neighbourhood of $p_0$, if it is nondegenerate at $p_0$.

**Compact-open topology.** The space of diffeomorphisms $\varphi$ is equipped with the *compact-open topology* on the complex analytic extensions of our real analytic systems (which is described in Section 3).

**Normal conjugacy.** Let $M_1$ be a manifold, and let $V \subset M_1$ be a submanifold of $M_1$. Let $T(M_1)$ and $T(V)$ denote the tangent bundles of $M_1$ and $V$ respectively, and let $T_V(M_1)$ be the restriction of $T(M_1)$ to $V$. The *normal bundle* $N(V)$ is defined as the quotient:

$$N(V) = T_V(M_1)/T(V).$$

Let $f : M_1 \to M_2$ be an embedding of the manifold $M_1$ in a manifold $M_2$: it induces a map $N_V(f) : N(V) \to N(f(V))$.

Now, let $\Phi, \varphi, \tilde{\varphi} : M \to M$ be diffeomorphisms, and let $V, W \subset M$ be submanifolds of $M$, such that $\varphi(V) = V$, $\Phi(V) = W$ and $\tilde{\varphi}(W) = W$. We say that $\Phi$ *conjugates* $\varphi$ *to* $\tilde{\varphi}$ *at the invariant manifold* $V$ if:

$$\Phi \Big|_V \circ \varphi \Big|_V = \tilde{\varphi} \Big|_W \circ \Phi \Big|_V .$$

Moreover, we say that $\Phi$ *normally conjugates* $\varphi$ *to* $\tilde{\varphi}$ *at the invariant manifold* $V$ if:

$$N_V(\Phi) \circ N_V(\varphi) = N_W(\tilde{\varphi}) \circ N_V(\Phi).$$

*2.2. Formulation of the result*

This subsection formulates the perturbation theorem. The proof is given in Sections 3 and 4.

**Theorem.** *Let* $1 \geqq \gamma > 0$ *and* $\tau > 2$ *be fixed. Suppose* $\varphi$ *is nondegenerate at* $p_0 \in P$*, with* $b(p_0) = (1, 0)$*,* $m(p_0) = 0$ *and* $\ell(p_0) = 0$*, and* $\mathcal{E}$ *is the manifold of codimension* $6k$ *through* $p_0$ *given by* $\mathcal{E} = \{p \in P \mid m(p) = 0 \text{ and } \ell(p) = 0\}$*.*

*Then there exists a neighbourhood* $\mathcal{N}$ *of* $p_0$ *in* $\mathcal{E}$*, and a neighbourhood* $\mathcal{V}$ *of* $\varphi$ *in the class of "vertical" diffeomorphisms on* $X \times P$*, such that for all small* $\tilde{\varphi} \in \mathcal{V}$ *there is a map* $\Phi : X \times \mathcal{N} \to X \times P$ *with the following properties:*

(a) *The map* $\Phi$ *is a* $C^\infty$ *diffeomorphism onto its image and lies in a small* $C^\infty$ *neighbourhood of the identity. Moreover,* $\Phi$ *is of the form*

$$\Phi(x, y, p) = \big(\Phi_p(x, y), \Pi(p)\big),$$

*where* $\Phi_p : X \times \mathcal{N} \to X$ *and* $\Pi : \mathcal{N} \to P$*; that is,* $\Phi$ *preserves the projection to* $P$*. The map* $\Phi_p$ *is affine (equal to its normal linear part) in* $y$*, and is real analytic in* $(x, y)$*.*

(b) *For every* $p \in \mathcal{N}_c$*, the diffeomorphism* $\Phi_p$ *normally conjugates* $\varphi_p$ *to* $\tilde{\varphi}_{\Pi(p)}$ *at the invariant circle* $S^1 \times \{0\} \subset X$*.*

**Remark.** The theorem states that the following diagram commutes for every $p \in \mathcal{N}_c$:

$$\begin{array}{ccc}
N\big(S^1 \times \{0\}\big) & \xrightarrow{N(\varphi_p)} & N\big(S^1 \times \{0\}\big) \\
\Big\downarrow {\scriptstyle N(\Phi_p)} & & \Big\downarrow {\scriptstyle N(\Phi_p)} \\
N\big(\Phi_p(S^1 \times \{0\})\big) & \xrightarrow{N(\tilde{\varphi}_{\Pi(p)})} & N\big(\Phi_p(S^1 \times \{0\})\big).
\end{array} \tag{10}$$

The diffeomorphism $\Phi_p$ will be a (normal) conjugacy only if the parameter $p$ is restricted to the parameter set $\mathcal{N}_c$. Outside this set, most likely the conjugating property will be lost.

Another, more intuitive formulation of the theorem is the following: the conjugacy $\Phi$ is of the form:

$$\Phi(x, y, p) = (x + u(x, p), y + v_0(x, p) + v_1(x, p)y, p + W(p)).$$

The map $\Pi : p \to p + W(p)$ shifts the set $\mathcal{N}_c$ a little in parameter space. Let $\tilde{\mathcal{N}}_c$ be its image (sketched in Fig. 1). The theorem states that *if* the parameter $p$ is such that $\tilde{p} = \Pi(p) = p + W(p) \in \tilde{\mathcal{N}}_c$, the diffeomorphism

$$\begin{aligned}
\tilde{\varphi}(x, y, \tilde{p}) = \Big(&x + a(\tilde{p}) + g(x, y, \tilde{p}),\\
&E(x, b(\tilde{p}))y + M(x, m(\tilde{p})) + L(x, \ell(\tilde{p}))y + g(x, y, \tilde{p})\Big),
\end{aligned}$$

can be conjugated (normally at $S \times \{0\}$) to

$$\varphi(x, y, p) = \Big(x + a(p) + \chi(x, y, p), E(x, b(p))y + \psi(x, y, p)\Big) \tag{11}$$

for some $\chi(x, y, p) = \mathrm{O}(|y|)$ and $\psi(x, y, p) = \mathrm{O}(|y|^2)$. Since $\varphi$ possesses an invariant circle with quasi-periodic dynamics, $\tilde{\varphi}$ does as well. Moreover, the normal behaviour of $\tilde{\varphi}$ equals that of $\varphi$.

Remark that the conjugacy $\Phi$, and thus the form of the set $\tilde{\mathcal{N}}_c$, depend both on $\tilde{\varphi}$.

## 3. Proof of the persistence theorem

This section, together with the next, gives the proof of the theorem stated in Section 2. Here, as a preliminary to the proof, the theorem is restated in a more technical (and analytical) form. Then a couple of propositions are given, which together prove the theorem. The next section will provide proofs of those propositions. Although different in its techniques and details, this proof is build on a framework, and contains many ideas, that can be found in [2], [3], [4], [11], [13].

### *3.1. Preliminary remarks*

In this subsection preparations are made in order to reformulate the main theorem. Recall that we are considering the "unperturbed" vertical diffeomorphism:

$$\begin{aligned}\varphi(x, y, p) = (&x + a(p) + \chi(x, y, p),\\ &E(x, b(p))y + M(x, m(p)) + L(x, \ell(p))y + \psi(x, y, p), p)\,.\end{aligned}$$

Here $x \in S^1$, $y \in \mathbf{R}^2$, $p \in P$, where $P \subset \mathbf{R}^q$ is a neighbourhood of a point $p_0$. The functions $a$, $b$, $m$ and $\ell$ are real analytic, and take values: $a(p) \in \mathbf{R}$, $b(p) \in \mathbf{R}^2$, $m(p) \in \mathbf{R}^{2k}$ and $\ell(p) \in \mathbf{R}^{4k}$. By assumption of the theorem, we have at $p_0$ that $a(p_0)$ is in the set of Diophantine frequencies $\mathbf{R}_c$; $b(p_0) = (1, 0)$, $m(p_0) = 0$ and $\ell(p_0) = 0$. The maps $E$, $M$ and $L$ are defined in (6)–(8) respectively. Finally we have that $\chi(x, y, p) = \mathrm{O}(|y|)$ and $\psi(x, y, p) = \mathrm{O}(|y|^2)$.

**3.1.1. Reparametrization.** Since $\varphi$ is assumed to be nondegenerate at $p_0$ (see Section 2.1), by the Inverse Function Theorem, there exists near $X \times \{p_0\}$ a reparametrization

$$V : (x, y, p) \mapsto (x, y, a(p), b(p), m(p), \ell(p), \nu(p))$$

(i.e., there exists a function $\nu(p)$ such that $V$ is a diffeomorphism), that conjugates $\varphi$ to $\tilde{\varphi}$; defining the new functions $\tilde{\chi}$ and $\tilde{\psi}$:

$$\begin{aligned}\tilde{\varphi}(x, y, a, b, m, \ell, \nu) &= V \circ \varphi \circ V^{-1} \\ &= \Big(x + a + \tilde{\chi}(x, y, a, b, m, \ell, \nu), \\ & E(x, b)y + M(x, m) + L(x, \ell)y + \tilde{\psi}(x, y, a, b, m, \ell, \nu), a, b, m, \ell, \nu\Big).\end{aligned}$$

Set $a_0 = a(p_0)$, $b_0 = b(p_0)[= (1, 0)]$, $m_0 = m(p_0)[= 0]$ and $\ell_0 = \ell(p_0)[= 0]$. From now on the parameter $\nu$ is dropped: it turns out that it can be easily incorporated again. Also the tilde on $\tilde{\varphi}$, $\tilde{\chi}$ and $\tilde{\psi}$ will be dropped. Thus $P$ is replaced by an open piece of $\mathbf{R}^{6k+3}$.

Below, the letter $p$ will be re-defined as $p = (a, b, m, \ell)$.

**3.1.2. Real analytic neighbourhoods.** Here the neighbourhood $\mathcal{V}$ of the "unperturbed" vertical diffeomorphism $\varphi$, as mentioned in the theorem of Section 2, will be described, as well as the compact-open topology of (real) analytic functions.

The function $f : U \to \mathbf{R}$, defined on an open subset $U$ of $\mathbf{R}$, is said to be *real analytic*, if there is an open complex neigbourhood $V \subset \mathbf{C}$ of $U$, and a complex analytic function $\tilde{f} : V \to \mathbf{C}$, such that $\tilde{f}$ restricted to $U$ equals $f$.

A remark on norms: for vectors $z \in \mathbf{R}^n$ or $\mathbf{C}^n$, $|z|$ will denote the maximum norm,

$$|z| = \max_i |z_i|.$$

As above, $\|b\|$ will denote the Euclidean norm (usually for 2-vectors),

$$\|b\| = \sqrt{b_1^2 + b_2^2}.$$

Functions $h : D \to \mathbf{C}^n$ will be equipped with the sup-norm,

$$|h|_D = \sup_D |h(z)|.$$

If a set $A \subset \mathbf{R}^n$ is considered, then the set $A + \varepsilon$ will be its $\varepsilon$-neighbourhood in the complex plane:

$$A + \varepsilon = \{z \in \mathbf{C}^n \;\Big|\; \inf_{w \in A} |z - w| < \varepsilon\}.$$

Let $\mathcal{A}$, $\mathcal{M}$ and $\mathcal{L}$ be compact neighbourhoods of respectively $a_0$, $m_0$ and $\ell_0$ in respectively $\mathbf{R}$, $\mathbf{R}^{2k}$ and $\mathbf{R}^{4k}$. Let $\mathcal{B}$ be the set

$$\left\{b \in \mathbf{R}^2 \,:\, \Big|\|b\| - 1\Big| \leqq k\frac{r_0}{8}\right\}$$

for some $0 < r_0 < 1$. Define

$$\mathcal{P} = \mathcal{A} \times \mathcal{B} \times \mathcal{M} \times \mathcal{L}.$$

Let $p = (a, b, m, \ell)$, as announced above. Finally, let $U$ be a neighbourhood of 0 in $\mathbf{R}^2$.

Let $O$ be some compact neighbourhood of $S^1 \times U \times \mathcal{P}$ in $(\mathbf{C}/(2\pi\mathbf{Z})) \times \mathbf{C}^2 \times \mathbf{C}^{6k+3}$. Without loss of generality $O$ can be taken to be of the form

$$O = \overline{\left(S^1 + r_0\right) \times (U + s_0) \times (\mathcal{A} + \rho_0) \times (\mathcal{B} + \alpha_0) \times (\mathcal{M} + s_0) \times (\mathcal{L} + q_0)},$$

with constants $0 < s_0, \rho_0, \alpha_0, q_0 \leqq 1$. For sufficiently small values of these constants, the diffeomorphism $\varphi$ has a complex analytic extension to such a set $O$.

Now $\mathcal{V}$ is determined by $O$, $\gamma$ and a positive constant $\delta_0$ given in the proof below. The set $\mathcal{V}$, mentioned in the theorem, consists of all families of diffeomorphisms $\tilde{\varphi} = \tilde{\varphi}(x, y, p)$ of the form

$$\begin{aligned}&\tilde{\varphi}(x, y, p) = \\ &\qquad (x + a + f(x, y, p), E(x, b)y + M(x, m) + L(x, \ell)y + g(x, y, p), p)\end{aligned}$$

with real analytic $f$ and $g$, both defined on $O$, such that in the supremum norm $|\cdot|_O$ on $O$ one has

$$|f|_O + \frac{|g|_O}{s_0} > \delta_0.$$

## 3.2. The Theorem

This subsection translates the geometrical formulation of the theorem of Section 2 into a more analytical one. Then an indication is given of the work to be done.

**3.2.1. Technical reformulation.** The theorem is restated in the following

**Reformulation.** *There is a constant $\delta_0 > 0$, depending on $\gamma$ and $O$, such that for all $f$ and $g$ satisfying:*

$$|f|_O + \frac{|g|_O}{s_0} < \delta_0,$$

*the following holds.*

*There are $C^\infty$ functions $\tilde{a}$, $\tilde{b}$, $\tilde{m}$ and $\tilde{\ell}$ from:*

$$\mathcal{N} \subset \mathcal{E} = \{p \in P \;\big|\; p = (a, b, 0, 0) \in \mathbf{R} \times \mathbf{R}^2 \times \mathbf{R}^{2k} \times \mathbf{R}^{4k}\}$$

*to $\mathbf{R}$, $\mathbf{R}^2$, $\mathbf{R}^{2k}$ and $\mathbf{R}^{4k}$ respectively, denoted collectively by*

$$\tilde{p}(p) = (\tilde{a}(p), \tilde{b}(p), \tilde{m}(p), \tilde{\ell}(p)),$$

*and real analytic functions $\chi(x, y, p)$ and $\psi(x, y, p)$ which are $\mathrm{O}(|y|)$ and $\mathrm{O}(|y|^2)$ respectively, with the following property. If $p$ is restricted to $\mathcal{N}_c$, that is,*

$$p = (a, b, 0, 0) \quad \textit{with} \quad a \in \mathcal{A}_c \textit{ and } b \in \mathcal{B},$$

*then the map*

$$\begin{aligned}\tilde{\varphi}(x, y, \tilde{p}(p)) &= (x + \tilde{a}(p) + f(x, y, \tilde{p}(p)), \\ &\qquad E(x, \tilde{b}(p))y + M(x, \tilde{m}(p)) + L(x, \tilde{\ell}(p))y + g(x, y, \tilde{p}(p)), \tilde{p}(p)\Big)\end{aligned}$$

*is conjugated to*

$$\varphi(x, y, p) = (x + a + \chi(x, y, p), E(x, b)y + \psi(x, y, p), p)\,.$$

*Moreover the conjugacy $\Phi$, satisfying $\Phi \circ \varphi = \tilde{\varphi} \circ \Phi$ for $p \in \mathcal{N}_c$, can be chosen to be the restriction to $X \times \mathcal{N}_c$ of a $C^\infty$ diffeomorphism $\Phi : X \times \mathcal{N} \to X \times P$, which is real analytic in $x$ and affine in $y$.*

The domain of definition of the map $\Phi$ in the parameter direction has to be slightly modified, in order to avoid problems at the boundary of the present set $\mathcal{P}$. We have to restrict to

$$\mathcal{P}' = \{p \in \mathcal{P} \mid \operatorname{dist}(p, \partial\mathcal{P}) > c_0\delta_0\}$$

for some $c_0$, depending only on $\gamma, \tau, k$ and $r_0$. Here $\partial\mathcal{P}$ denotes the boundary of $\mathcal{P}$. For $\delta_0 > 0$ sufficiently small, this is still a neighbourhood. Also denote $\mathcal{N} \cap \mathcal{P}'$ by $\mathcal{N}'$ etc.

**3.2.2. Idea of the proof.** Let us briefly indicate the idea of the proof. Recall that a map $\Phi : X \times \mathcal{N}_c' \to X \times P$ has to be found, preserving the projection to $P$, which conjugates $\varphi$ to $\tilde{\varphi}$, i.e., such that

$$\Phi \circ \varphi = \tilde{\varphi} \circ \Phi. \tag{12}$$

Write

$$\Phi(x, y, p) = \big(x + \bar{u}(x, p), y + \bar{v}(x, y, p), p + \bar{W}(p)\big),$$

where

$$\bar{v}(x, y, p) = \bar{v}_0(x, p) + \bar{v}_1(x, p)y$$

and

$$\bar{W}(p) = (\bar{w}_1(p), \bar{w}_2(p), \bar{w}_3(p), \bar{w}_4(p)).$$

Note that $p + \bar{W}(p) = \tilde{p}(p)$. Furthermore

$$\varphi(x, y, p) = \Big(x + a + \chi(x, y, p), E(x, b)y + \psi(x, y, p), p\Big)$$

if $p = (a, b, 0, 0)$, and

$$\tilde{\varphi}(x, y, p) = \Big(x+a+f(x, y, p), E(x, b)y+M(x, m)+L(x, \ell)y+g(x, y, p), p\Big)$$

for general $p \in P$.

Writing (12) out componentwise yields

$$\begin{aligned} x + a + \chi(x, y, p) &+ \bar{u}(x + a + \chi, p) \\ &= x + \bar{u}(x, p) + a + \bar{w}_1(p) + f(x + \bar{u}, y + \bar{v}, p + \bar{W}) \end{aligned} \tag{13}$$

and

$$\begin{aligned} E(x, b)y + \psi(x, y, p) &+ \bar{v}\Big(x + a + \chi, E(x, b)y + \psi, p\Big) \\ &= E\Big(x + \bar{u}, b + \bar{w}_2(p)\Big)(y + \bar{v}) + g\Big(x + \bar{u}, y + \bar{v}, p + \bar{W}\Big) \\ &\quad + M\Big(x + \bar{u}, m + \bar{w}_3(p)\Big) + L\Big(x + \bar{u}, \ell + \bar{w}_4(p)\Big)(y + \bar{v}). \end{aligned} \tag{14}$$

These are complicated nonlinear equations in $\bar{u}$, $\bar{v}$, $\bar{W}$, $\chi$ and $\psi$. They are solved iteratively by a Newtonian procedure. That is, the equations will be approximated by equations, which are almost linearizations of (13) and (14), but not quite. They furnish approximations to the functions $\bar{u}$, $\bar{v}$ and $\bar{W}$, determining the diffeomorphism $\Phi$. The approximating equations are either of the well-known form (see [1])

$$u(x+a,p) - u(x,p) = f(x,0,p) + w_1(p)$$

(which involves a small divisor problem), or of the form (see [6])

$$v(x+a,0,p) - E(x,b)v(x,0,p) = g(x,0,p) + M(x,w_3(p))$$

(which is not a small divisor problem at all). Having solved the approximating equations, functions $\chi$ and $\psi$ are then determined by requiring that relation (12) holds on $X \times \mathcal{E}$. Because the conjugacy equations have been approximated, the functions $\chi$ and $\psi$ are not of order $\mathrm{O}(|y|)$ and $\mathrm{O}(|y|^2)$; however, they are much smaller than $f$ and $g$ respectively, and so they provide the starting point of another iteration step. As it stands, this reasoning is largely heuristical; it is made more precise in the next subsection.

### *3.3. Proof*

The proof of the theorem is divided into three parts. First the framework of the Newtonian induction process is presented in Sections 3.3.1–3.3.5. Secondly the estimates needed in the iteration step are given in Section 3.3.6, and finally, in Section 3.3.7, a proposition is stated which implies convergence of the process.

**3.3.1. Frame of the proof.** Given the perturbation $\tilde{\varphi}$ of $\varphi$, the conjugacy $\Phi$ solving (13) and (14) will be obtained as a Whitney-$C^\infty$ limit of a sequence $\{\Phi_j\}_{j=0}^{\infty}$ of (real) analytic diffeomorphisms, defined on complex neighbourhoods $\mathcal{D}_j$ of $X \times \mathcal{N}_c'$. Here the inverse approximation lemma will be used ([4]). See Appendix A for a statement of this lemma. In order to describe the inductive construction of the $\Phi_j$ the following notation is introduced.

Define $\tilde{\varphi}_j = \Phi_j \circ \tilde{\varphi} \circ \Phi_j^{-1}$. For $j = 0$ we put $\Phi_0 = \mathrm{id}$; then $\tilde{\varphi}_0 = \tilde{\varphi}$. For $j \geqq 1$, the maps $\Phi_j$ and $\tilde{\varphi}_j$ take the following forms ($x_j = x|_{\mathcal{D}_j}$ etc.):

$$\Phi_j(x_j,y_j,p_j) = \\ \Big(x_j + \bar{u}^j(x_j,p_j),\, y_j + \bar{v}_0^j(x_j,p_j) + \bar{v}_1^j(x_j,p_j)y_j,\, p_j + \bar{W}^j(p_j)\Big)$$

and

$$\tilde{\varphi}_j(x_j,y_j,p_j) = \Big(x_j + a_j + f^j(x_j,y_j,p_j), \\ E(x_j,b_j)y_j + M(x_j,m_j) + L(x_j,\ell_j)y_j + g^j(x_j,y_j,p_j),\, p_j\Big).$$

In view of the inverse approximation lemma, care has to be taken that both $\Phi_j$ and $\tilde{\varphi}_j$ do have analytic extensions to a complex neighbourhood $\mathcal{D}_j$ of the Cantor set $S^1 \times \{0\} \times \mathcal{N}_c'$. The neighbourhoods will satisfy $\mathcal{D}_j \subset O$ and will shrink, for $j \to \infty$, in an appropriate (geometrical) way in the $a$ and $b$ directions, and supergeometrically in the $\ell$ and $m$ directions. They will be specified precisely in Section 3.3.5.

During the discussion of the induction steps, it will be established that

$$|f^j|_{\mathcal{D}_j},\ |g^j|_{\mathcal{D}_j}/s_j \to 0$$

very fast (faster than any geometrical sequence). Then application of the inverse approximation lemma gives limits $\Phi_\infty$ and $\tilde{\varphi}_\infty$ of the sequences $\{\Phi_j\}$ and $\{\tilde{\varphi}_j\}$ respectively, which are Whitney-$C^\infty$ on the closed set $S^1 \times \{0\} \times \mathcal{N}'_c$. The desired map $\Phi$, defined on $S^1 \times U \times \mathcal{N}$, finally will be obtained by the Whitney extension theorem ([16]), applied to $\Phi_\infty$.

In the coordinates $(x_\infty, y_\infty, p_\infty)$ the diffeomorphism $\tilde{\varphi}_\infty = \Phi_\infty \circ \tilde{\varphi} \circ \Phi_\infty^{-1}$ will have the form

$$\begin{aligned}&\tilde{\varphi}_\infty(x_\infty, y_\infty, p_\infty) = \\ &\big(x_\infty + a_\infty + f^\infty(x_\infty, y_\infty, p_\infty), E(x_\infty, b_\infty)y_\infty + g^\infty(x_\infty, y_\infty, p_\infty), p_\infty\big),\end{aligned}$$

where

$$f^\infty(x_\infty, y_\infty, p_\infty) = \mathrm{O}(|y_\infty|) \quad \text{and} \quad g^\infty(x_\infty, y_\infty, p_\infty) = \mathrm{O}(|y_\infty|^2);$$

this proves the theorem. In the following, the first step is to establish the properties of one general induction step; then the inverse approximation lemma can be applied as sketched above.

**3.3.2. Description of the induction.** Suppose a sequence $\{\Phi_i\}_{i=0}^j$ as described above has already been obtained. In order to see what happens in the induction step, write $\Phi_{j+1} = \Phi_j \circ \mathcal{U}_j$ for maps

$$\mathcal{U}_j : \mathcal{D}_{j+1} \to \mathcal{D}_j, \quad j \geqq 0.$$

Then for all $j \geqq 1$,

$$\Phi_j = \mathcal{U}_0 \circ \mathcal{U}_1 \circ \cdots \circ \mathcal{U}_{j-1} \qquad \text{and} \qquad \tilde{\varphi}_{j+1} = \mathcal{U}_j^{-1} \circ \tilde{\varphi}_j \circ \mathcal{U}_j.$$

This can be illustrated by the following commuting diagram:

$$\begin{tikzcd}
\mathcal{D}_0 \arrow[d, "\varphi_0"] & \mathcal{D}_1 \arrow[l, "\mathcal{U}_0"'] \arrow[d, "\varphi_1"] & \cdots \arrow[l, "\mathcal{U}_1"'] & \mathcal{D}_j \arrow[l, "\mathcal{U}_{j-1}"'] \arrow[d, "\varphi_j"] & \mathcal{D}_{j+1} \arrow[l, "\mathcal{U}_j"'] \arrow[d, "\varphi_{j+1}"] \\
\varphi_0(\mathcal{D}_0) & \varphi_1(\mathcal{D}_1) \arrow[l, "\mathcal{U}_0"'] & \cdots \arrow[l, "\mathcal{U}_1"'] & \varphi_j(\mathcal{D}_j) \arrow[l, "\mathcal{U}_{j-1}"'] & \varphi_{j+1}(\mathcal{D}_{j+1}) \arrow[l, "\mathcal{U}_j"']
\end{tikzcd} \tag{15}$$

Note that since $\Phi_0 = \mathrm{id}$, the diffeomorphism $\tilde{\varphi}_0$ is the original diffeomorphism $\tilde{\varphi}$. In the following, the relationship between $\tilde{\varphi}_j$ and $\tilde{\varphi}_{j+1}$ will be examined.

As long as a general iteration step is treated, the so-called $+$-notation will be used:

$$(x, y, a, b, m, \ell) \text{ is written instead of } (x_j, y_j, a_j, b_j, m_j, \ell_j),$$

and

$$(\xi, \eta, \omega, \beta, \mu, \lambda) \text{ instead of } (x_{j+1}, y_{j+1}, a_{j+1}, b_{j+1}, m_{j+1}, \ell_{j+1}).$$

Furthermore, $p$ is written instead of $p_j$ and $\sigma$ instead of $p_{j+1}$; also $f^j$ will be replaced by $f$ and $f^{j+1}$ by $f^+$, $\mathcal{D}_j$ by $\mathcal{D}$ and $\mathcal{D}_{j+1}$ by $\mathcal{D}_+$, etc. Note that Greek letters are used to denote variables in the domain $\mathcal{D}_+$, while Latin letters are used for variables in the domain $\mathcal{D}_j$. For instance, in the new notation the diffeomorphism $\tilde{\varphi}_+$ has the form

$$\tilde{\varphi}_+(\xi, \eta, \sigma) = \\ \big(\xi + \omega + f^+(\xi, \eta, \sigma), E(\xi, \beta)\eta + M(\xi, \mu) + L(\xi, \lambda)\eta + g^+(\xi, \eta, \sigma), \sigma\big).$$

The map $\mathcal{U} : \mathcal{D}_+ \to \mathcal{D}$, whenever defined, will be taken to be of the form

$$\mathcal{U}(\xi, \eta, \omega, \beta, \mu, \lambda) = \begin{pmatrix} \xi + u(\xi, \sigma) \\ \eta + v(\xi, \eta, \sigma) \\ \omega + w_1(\sigma) \\ \beta + w_2(\sigma) \\ \mu + w_3(\sigma) \\ \lambda + w_4(\sigma) \end{pmatrix}, \tag{16}$$

where $v(\xi, \eta, \sigma) = v_0(\xi, \sigma) + v_1(\xi, \sigma)\eta$. The relation $\mathcal{U} \circ \tilde{\varphi}_+ = \tilde{\varphi} \circ \mathcal{U}$ then reads

$$\begin{aligned} &\xi + \omega + f^+(\xi, \eta, \sigma) + u(\xi + \omega + f^+, \sigma) \\ &\qquad = \xi + u(\xi, \sigma) + \omega + w_1(\sigma) + f\Big(\xi + u, \eta + v, \sigma + W\Big) \end{aligned} \tag{17}$$

and

$$\begin{aligned} &E(\xi, \beta)\eta + g^+(\xi, \eta, \sigma) + M(\xi, \mu) + L(\xi, \lambda)\eta \\ &\quad + v\Big(\xi + \omega + f^+, E(\xi, \beta)\eta + g^+ + M(\xi, \mu) + L(\xi, \lambda)\eta, \sigma + W(\sigma)\Big) \\ &\quad = E\Big(\xi + u, \beta + w_2(\sigma)\Big)\Big(\eta + v(\xi, \eta, \sigma)\Big) + g\Big(\xi + u, \eta + v, \sigma + W\Big) \\ &\qquad\qquad + M\Big(\xi + u, \mu + w_3(\sigma)\Big) + L\Big(\xi + u, \lambda + w_4(\sigma)\Big)(\eta + v) \end{aligned} \tag{18}$$

for all $(\xi, \eta, \sigma) \in \mathcal{D}_+$. Compare (17) and (18) with (13) and (14): the main difference is that here $\mu$ and $\lambda$ (and consequently $M(\xi, \mu)$ and $L(\xi, \lambda)$) are not zero in $\tilde{\varphi}_+$.

The aim is to choose $\mathcal{U}$ such that $f^+$ and $g^+$ are much smaller (in a precise sense) than $f$ and $g$. Note that this still leaves us some freedom in the choice of $u$ and $v$. As remarked above, in the limit $j \to \infty$, the functions $f^j$ and $g^j$ tend to 0 very fast.

**3.3.3. Determination of the conjugacy.** Here the map $\mathcal{U} : \mathcal{D}_+ \to \mathcal{D}$, see (16), is determined, which conjugates $\tilde{\varphi}$ to $\tilde{\varphi}_+$, determining $\tilde{\varphi}_+$. In order to make the Newtonian iteration work, the function $u$ (see the first component of $\mathcal{U}$), and the first component $w_1$ of $W$, are required to be a solution of the following approximation of (17):

$$u(\xi + \omega, \sigma) - u(\xi, \sigma) = w_1(\sigma) + {}_d f(\xi, 0, \sigma). \tag{19}$$

Here ${}_d f$, for $d = d_j \in \mathbf{N}$, is a truncation of the Fourier series of $f$ at the order $d$. That is, if $f = \sum f_n e^{in\xi}$, then

$$ {}_d f = \sum_{|n| \leqq d} f_n e^{in\xi}. $$

The sequence $\{d_j\}$ is chosen later on in an appropriate way (in Section 3.3.7). Equation (19) poses a small divisor problem [1]. It will be solved eventually in Section 4.2.1.

The function $v$, and the components $(w_2, w_3, w_4)$ of $W$ are required to form a solution of the following approximation of (18):

$$ \begin{aligned} & v(\xi+\omega, E(\xi,\beta)\eta, \sigma) - E(\xi,\beta)v(\xi,\eta,\sigma) \\ &\quad = E(\xi, w_2)\eta + M(\xi, w_3) + L(\xi, w_4)\eta \\ &\qquad + {}_d g(\xi, 0, \sigma) + {}_d g_\eta(\xi, 0, \sigma)\eta + {}_d\Big\{E(\xi+u,\beta) - E(\xi,\beta)\Big\}\eta \\ &\qquad + {}_d\Big\{M(\xi+u,\mu) - M(\xi,\mu)\Big\} + {}_d\Big\{L(\xi+u,\lambda) - L(\xi,\lambda)\Big\}\eta. \end{aligned} \tag{20} $$

Here it is assumed that $u(\xi, \sigma)$ is already determined from the first equation. Equation (20) splits into a part independent of $\eta$ and a part linear in $\eta$, which are solved separately:

$$ \begin{aligned} & v_0(\xi+\omega, \sigma) - E(\xi,\beta)v_0(\xi,\sigma) \\ &\quad = M(\xi, w_3) + {}_d g(\xi, 0, \sigma) + {}_d\Big\{M(\xi+u,\mu) - M(\xi,\mu)\Big\} \end{aligned} \tag{21} $$

and

$$ \begin{aligned} & v_1(\xi+\omega, \sigma)E(\xi,\beta) - E(\xi,\beta)v_1(\xi,\sigma) \\ &\quad = E(\xi, w_2) + L(\xi, w_4) + {}_d g_\eta(\xi, 0, \sigma) \\ &\qquad + {}_d\Big\{E(\xi+u,\beta) - E(\xi,\beta) + L(\xi+u,\lambda) - L(\xi,\lambda)\Big\}. \end{aligned} \tag{22} $$

In this subsection, various choices have been made: which term to use in these equations, which terms to leave for the remainder equations, etc. These choices have been made in such a way that the induction will work; that is, so that the "remainder terms" $f^+$ and $g^+$ will be sufficiently small. Informally speaking, this is achieved by putting "as many terms as possible" into the approximations; the process is emphatically non-unique.

**3.3.4. Determination.** Expressions for the "remainder terms" $f^+$ and $g^+$ are obtained by subtracting (19) from (17), and (20) from equation (18). We get for $f^+$:

$$ \begin{aligned} & \xi + \omega + f^+(\xi,\eta,\sigma) + u(\xi+\omega+f^+,\sigma) - u(\xi+\omega,\sigma) \\ &\quad = \xi + \omega + f(\xi+u, \eta+v, \sigma+W) - {}_d f(\xi, 0, \sigma). \end{aligned} $$

Re-written, this reads:

$$ f^+(\xi,\eta,\sigma) + u(\xi+\omega+f^+,\sigma) - u(\xi+\omega,\sigma) = t_1 + t_2 + t_3, \tag{23} $$

where

$$
\begin{aligned}
t_1 &= f(\xi+u, \eta+v, \sigma+W) - f(\xi, \eta, \sigma),\\
t_2 &= f(\xi, \eta, \sigma) - f(\xi, 0, \sigma),\\
t_3 &= f(\xi, 0, \sigma) - {}_d f(\xi, 0, \sigma).
\end{aligned}
$$

Likewise for $g^+$ the following expression is obtained:

$$
g^+(\xi, \eta, \sigma) + v_1(\xi+\omega+f^+, \sigma+W)g^+ = \sum_{i=4}^{17} t_i, \tag{24}
$$

where (remembering that $M(\xi, \mu)$ and $L(\xi, \lambda)$ are linear in $\mu$ and $\lambda$ respectively)

$$
\begin{aligned}
t_4 &= g(\xi+u, \eta+v, \sigma+W) - g(\xi, \eta, \sigma),\\
t_5 &= g(\xi, \eta, \sigma) - g(\xi, 0, \sigma) - g_\eta(\xi, 0, \sigma)\eta,\\
t_6 &= g(\xi, 0, \sigma) + g_\eta(\xi, 0, \sigma)\eta - {}_d\left\{g(\xi, 0, \sigma) - g_\eta(\xi, 0, \sigma)\eta\right\},\\
t_7 &= M(\xi+u, w_3) - M(\xi, w_3),\\
t_8 &= \{E(\xi+u, w_2) - E(\xi, w_2)\}\,\eta,\\
t_9 &= \{E(\xi+u, \beta+w_2) - E(\xi, \beta)\}\,v(\xi, \eta, \sigma),\\
t_{10} &= \{L(\xi+u, w_4) - L(\xi, w_4)\}\,\eta,\\
t_{11} &= \{L(\xi+u, \lambda+w_4) - L(\xi, \lambda)\}\,v(\xi, \eta, \sigma),\\
t_{12} &= L(\xi, \lambda)v(\xi, \eta, \sigma)\\
t_{13} &= M(\xi+u, \mu) - M(\xi, \mu) - {}_d\{M(\xi+u, \mu) - M(\xi, \mu)\},\\
t_{14} &= \{E(\xi+u, \beta) - E(\xi, \beta)\}\,\eta - {}_d\{E(\xi+u, \beta) - E(\xi, \beta)\}\,\eta,\\
t_{15} &= \{L(\xi+u, \lambda) - L(\xi, \lambda)\}\,\eta - {}_d\{L(\xi+u, \lambda) - L(\xi, \lambda)\}\,\eta,\\
t_{16} &= v\left(\xi+\omega+f^+, E(\xi, \beta) + M(\xi, \mu) + L(\xi, \lambda)\eta, \sigma+W(\sigma)\right)\\
&\quad - v\left(\xi+\omega, E(\xi, \beta) + M(\xi, \mu) + L(\xi, \lambda)\eta, \sigma+W(\sigma)\right),\\
t_{17} &= v_1(\xi+\omega, \sigma+W(\sigma))(M(\xi, \mu) + L(\xi, \lambda)\eta).
\end{aligned}
$$

Below (in Section 3.3.6) it will be established that $v_1$ and $\partial u/\partial\xi$ are small, so that $f^+$ and $g^+$ can be solved (and estimated) from (23) and (24) by the Implicit Function Theorem.

**3.3.5. Specification of the complex domains.** In this subsubsection the complex neighbourhoods $\mathcal{D}_j$ will be specified, as well as the orders of truncation $d_j$. The induction hypothesis is that the estimate

$$
|f^j|_{\mathcal{D}_j} + \frac{|g^j|_{\mathcal{D}_j}}{s_j} \leqq \delta_j \tag{25}
$$

holds, where $\{s_j\}_{j=0}^\infty$ and $\{\delta_j\}_{j=0}^\infty$ are decreasing sequences of positive real numbers, yet to be determined. In $+$-notation, this reads:

$$
|f| + \frac{|g|}{s} \leqq \delta.
$$

In order to define $\mathcal{D}_j$, we will need the sequences $\{s_j\}_{j=0}^{\infty}$ and $\{\delta_j\}_{j=0}^{\infty}$ mentioned above, and a sequence $\{q_j\}_{j=0}^{\infty}$, also still to be determined. These three sequences will be converging to zero faster than any geometrical sequence. The sequence $\{\delta_j\}$ will be of the form

$$\delta_+ = \delta^{\kappa},$$

where $\kappa > 1$. This implies that $\delta_j = \delta_0^{\kappa^j}$. The other two will be expressed in powers of $\delta$:

$$\frac{s_+}{s} = \delta^{\zeta_1},$$
$$q = \delta^{\zeta_2},$$

where $\zeta_1, \zeta_2 \in \mathbf{R}$ are still undetermined, as well as $\kappa$. Also some geometrically decaying sequences will be needed: $\{r_j\}_{j=0}^{\infty}$, $\{\rho_j\}_{j=0}^{\infty}$ and $\{\alpha_j\}_{j=0}^{\infty}$.

The following choices will be made in order that the induction works: $r_j$ is defined as

$$r_j = \frac{r_0}{2}\left(1 + 2^{-j}\right). \tag{26}$$

Note that $r_j$ converges to $r_0/2$ as $j \to \infty$, not to 0. Here an additional requirement is placed on $r_0$:

$$r_0 < \frac{1}{k}.$$

(This is the same "twisting" constant $k$ as in (6).) This will be used in the proof of Proposition 2.

Recall the function Entier$(x)$, giving the largest integer smaller than or equal to $x$:

$$\mathrm{Entier}(x) = \max\{n \in \mathbf{Z} : n \leqq x\}.$$

The truncation $d_j$ is defined as

$$d_j = \mathrm{Entier}\left((2\kappa)^j \frac{4}{r_0} \log \delta_0^{-1}\right).$$

The sequences $\rho_j$ and $\alpha_j$ are given by:

$$\rho_j = \frac{1}{2}\gamma d_j^{-\tau},$$
$$\alpha_j = \frac{kr_0}{8}2^{-j}.$$

Here $\gamma$ and $\tau$ are the same as in (9). Throughout the rest of the proof, it will be assumed that the rapidly converging sequences will be smaller everywhere than the geometrically converging sequences, that is

$$\delta_j, s_j, q_j < r_j, \rho_j, \alpha_j, d_j^{-1} \qquad \text{for all } j.$$

Moreover, it will be especially assumed that

$$\delta_j < \frac{\left(r_j - (\frac{1}{4}r_{j+1} + \frac{3}{4}r_j)\right)^{\tau+3}}{c' 4^{j\tau}}, \tag{27}$$

where $c'$ is given by Lemma 6 (see Section 4.2.1). This can be effected by choosing $\delta_0$, $s_0$ and $q_0$ sufficiently small.

The domains $\mathcal{D}_j$ are defined using the above sequences (see Section 3.1.2 for the definitions of $\mathcal{A}$ and $\mathcal{B}$, and (9) for the definition of the subscript $_c$):

$$\mathcal{D}_j = \left(S^1 + r_j\right) \times \left\{y \in \mathbf{C}^2 \;\middle|\; |y| < s_j\right\} \times \left(\mathcal{A}'_c + \rho_j\right) \times \left(\mathcal{B} + \alpha_j\right) \\ \times \left\{(\mu, \lambda) \in \mathbf{C}^{2k} \times \mathbf{C}^{4k} \;\middle|\; |\mu| < s_j,\ |\lambda| < q_j\right\},$$

or in $+$-notation

$$\mathcal{D} = \left(S^1 + r\right) \times \left\{y \in \mathbf{C}^2 \;\middle|\; |y| < s\right\} \times \left(\mathcal{A}'_c + \rho\right) \times (\mathcal{B} + \alpha) \\ \times \left\{(\mu, \lambda) \in \mathbf{C}^{2k} \times \mathbf{C}^{4k} \;\middle|\; |\mu| < s,\ |\lambda| < q\right\}.$$

More generally, for a real parameter $t$, with $0 \leqq t \leqq 1$, we define the sequences $r_{j+t}$, $s_{j+t}$ etc. (in $+$-notation $r_t$, $s_t$ etc.) by

$$r_{j+t} = t r_{j+1} + (1 - t) r_j,$$

and the domains $\mathcal{D}_{j+t}$ (in $+$-notation $\mathcal{D}_t$):

$$\mathcal{D}_t = \left(S^1 + r_t\right) \times \left\{y \in \mathbf{C}^2 \;\middle|\; |y| < s_t\right\} \times \left(\mathcal{A}'_c + \rho_t\right) \times (\mathcal{B} + \alpha_t) \\ \times \left\{(\mu, \lambda) \in \mathbf{C}^{2k} \times \mathbf{C}^{4k} \;\middle|\; |\mu| < s_t,\ |\lambda| < q_t\right\}.$$

The supremum of a function $f$ over $\mathcal{D}_t$ will be denoted by

$$|f|_t = \sup_{\mathcal{D}_t} |f|.$$

**3.3.6. Estimates for the induction step.** As mentioned in the previous subsubsection, the induction hypothesis reads

$$|f| + \frac{|g|}{s} \leqq \delta. \tag{28}$$

It is one of the assumptions of the reformulated theorem (Section 3.2.1) that the hypothesis is satisfied for $j = 0$. Using (28), estimates for $|f^{j+1}|$, $|g^{j+1}|$ and $|\mathcal{U}_j|_{\mathcal{D}_j}$ in terms of $\delta$ can be obtained. Eventually (28) has to be proved for $j+1$ to complete the induction. The estimates are obtained in three propositions, which are stated here and which will be proved in Sections 4.2, 4.3 and 4.4 respectively.

The first proposition gives estimates for the conjugacy $\mathcal{U}$. It is proved in Section 4.2:

**Proposition 1.** *Let $f$ and $g$ be analytic on $\mathcal{D}$ and let* (28) *hold. Then* (19), (21) *and* (22) *can be solved for $u$, $v$ and $W$; these functions can be chosen such that on $\mathcal{D}_{\frac{1}{2}}$ the following estimates for $(u, v, W) = (u, v, w_1, \ldots, w_4)$ hold:*

$$|u|_{\frac{1}{2}} + \frac{|v|_{\frac{1}{2}}}{s} \leqq c_1 8^{j\tau} \delta$$

*and*

$$|w_1|_{\frac{1}{2}} + 4^{-j\tau}|w_2|_{\frac{1}{2}} + 4^{-j\tau}\frac{|w_3|_{\frac{1}{2}}}{s} + 4^{-j\tau}|w_4|_{\frac{1}{2}} \leqq c_1\delta$$

*where $c_1$ depends only on $\tau$, $\gamma$, $k$ and $r_0$.*

The following corollary follows immediately:

**Corollary 2.** *There is a constant $\Delta \in (0, 1)$, dependent only of $\tau, \gamma, k, r_0, c'$ and $\kappa$, but not on $j$, such that if $0 < \delta_0 < \Delta$, then:*

$$\mathcal{U}(\mathcal{D}_+) \subset \mathcal{D}_{\frac{3}{4}}.$$

The remainder terms $f^+$ and $g^+$ can be determined from (23) and (24) as follows.

**Proposition 3.** *Let* (28) *hold. Assume that* $0 < \delta_0 < \Delta$. *Then $f^+$ and $g^+$ can be estimated from* (23) *and* (24) *by*

$$|f^+|_+ \leqq 2\left(|t_1|_+ + |t_2|_+ + |t_3|_+\right), \quad |g^+|_+ \leqq 2\sum_{i=4}^{17}|t_i|_+,$$

*where*

$$|t_1|_+ + \frac{|t_4|_+}{s} \leqq c_2 8^{j\tau}\frac{\delta^2}{q}, \quad |t_2|_+ \leqq c_2\frac{s_+}{s}\delta, \quad \frac{|t_5|_+}{s} \leqq c_2\frac{s_+^2}{s^2}\delta,$$

$$s|t_3|_+ + |t_6|_+ + |t_{13}|_+ + |t_{14}|_+ + |t_{15}|_+ \leqq c_2 2^j \mathrm{e}^{-d(r-r_+)} s\delta,$$

$$|t_7|_+ + |t_8|_+ + |t_9|_+ + |t_{10}|_+ + |t_{11}|_+ \leqq c_2 8^{j\tau} s\delta^2,$$

$$|t_{12}|_+ + |t_{17}|_+ \leqq c_2 4^{j\tau} s\delta q_+, \quad |t_{16}|_+ \leqq c_2 8^{j\tau} s\delta|f^+|_+.$$

*The constant $c_2$ depends only on $\tau$, $\gamma$, $k$ and $r_0$.*

This proposition is proved in Section 4.3.

In order to complete the induction, the induction assumption for $j+1$ has to be proved. Moreover the conditions of the Inverse Approximation Lemma have to be checked.

The existence of suitable sequences $\delta$, $s$ and $q$ is formulated in the following proposition, which is proved in Section 4.4:

**Proposition 4.** *Let* (28) *hold. If moreover we choose*

$$\delta_+ = \delta^{\frac{6}{5}}, \quad q = \delta^{\frac{1}{2}} \text{ and } s_+ = \delta^{\frac{1}{4}}s$$

*and if $\delta_0$ is chosen small enough (independent of $j$), then*

$$|f^+|_+ + \frac{|g^+|_+}{s_+} \leqq \delta_+.$$

This proposition proves the induction hypothesis for $j+1$. Starting from $f^0$ and $g^0$, the coordinate transform $\mathcal{U}_0$ is determined, and together with it the functions $f^1$ and $g^1$ etc., *ad infinitum.*

**3.3.7. Convergence.** The diffeomorphism $\tilde\varphi_{j+1}$ and the coordinate transformations $\mathcal{U}_j$ are defined on domains $\mathcal{D}_j$ which shrink down to a domain $\mathcal{D}_\infty$ in various ways. Using the results of the previous subsection and the Inverse Approximation Lemma, the existence and differentiability properties of the limiting diffeomorphism,

$$\Phi_\infty = \lim_{j\to\infty} \Phi_j = \lim_{j\to\infty} \mathcal{U}_0 \circ \mathcal{U}_1 \circ \mathcal{U}_2 \circ \cdots \circ \mathcal{U}_{j-1},$$

can be derived. The result is stated in the following proposition; the proof is relegated to Section 4.5.

**Proposition 5.** *Let the assumptions of the theorem hold. Then the diffeomorphisms*

$$\Phi_j = \mathcal{U}_0 \circ \mathcal{U}_1 \circ \cdots \circ \mathcal{U}_{j-1}$$

*converge to the restriction to* $\mathcal{D}_\infty = \bigcap \mathcal{D}_j$ *of a diffeomorphism* $\Phi_\infty$*, which is defined on a neighbourhood of* $\mathcal{D}_\infty$*, and which has the following properties:*

(a) $\Phi_\infty$ *is a* $C^\infty$ *diffeomorphism from* $X \times \mathcal{N}$ *onto its image.*
(b) *For fixed parameter values* $p \in \mathcal{N}'_c$*, the diffeomorphism* $\Phi_\infty$ *is real analytic on* $X$*.*
(c) $\Phi_\infty$*, restricted to* $X \times \mathcal{N}'_c$*, conjugates* $\tilde\varphi$ *to* $\tilde\varphi_\infty$*, where* $\tilde\varphi_\infty$ *is of the form*

$$\tilde\varphi_\infty(x, y, p) = \Big(x + a + f^\infty(x, y, p),\ E(x, b)y + g^\infty(x, y, p),\ \ p\Big)$$

*with* $f = \mathrm{O}(|y|)$ *and* $g = \mathrm{O}(|y|^2)$.

This completes the proof of the theorem.

## 4. Proof of the reformulated persistence theorem

In this section, the propositions stated in the previous section are proved.

### *4.1. Preliminaries*

For notational convenience, in the following estimates the so-called dot-notation is introduced: constants depending only on the "uniform" (with respect to $j$) constants $\tau$, $\gamma$, $k$ and $r_0$ will sometimes be indicated by a dot following the inequality sign: $\ldots \leqq\cdot\ \ldots$.

Note for instance that from the definition (26) of $\{r_j\}_{j=0}^\infty$ in the previous section, it follows that

$$(r - r_+)^{-1}, (r - r_{\frac12})^{-1}, (r - r_{\frac14})^{-1} \leqq\cdot\ 2^j\ .$$

Remark that the following estimates on the $\frac{\partial}{\partial\xi}$-derivatives of the functions $E$, $M$ and $L$ hold (see for the definitions of $E$, $M$ and $L$ (6)–(8)):

$$\left|\frac{\partial E}{\partial\xi}(\xi, \beta)\right| \leqq\cdot\ |\beta|, \qquad \left|\frac{\partial M}{\partial\xi}(\xi, \mu)\right| \leqq\cdot\ |\mu|, \qquad \left|\frac{\partial L}{\partial\xi}(\xi, \lambda)\right| \leqq\cdot\ |\lambda|.$$

Below Fourier coefficients as well as derivatives of real analytic functions have to be estimated; for this the following two lemmas are needed. The first goes back to the Paley-Wiener theorem:

**Paley-Wiener Lemma.** *Let $h(\xi)$ be real analytic with complex extension to $S^1+\zeta$. Let*

$$\sum_{n=-\infty}^{\infty} h_n \mathrm{e}^{in\xi}$$

*be its Fourier series. Then the coefficients $h_n$ can be estimated:*

$$|h_n| \leqq 2\pi |h|_\zeta \mathrm{e}^{-|n|\zeta}$$

*where $|h|_\zeta$ is the supremum of $|h|$ over $S^1+\zeta$.*

*Conversely, if coefficients $|h_n|$ satisfy an estimate of the form*

$$|h_n| \leqq K \mathrm{e}^{-|n|\zeta},$$

*then, for $0 < \zeta_+ < \zeta < 1$, the function $h(\xi) = \sum h_n \mathrm{e}^{in\xi}$ is analytic on $S^1+\zeta_+$ and can be estimated by*

$$|h|_{\zeta_+} \leqq c \frac{K}{\zeta - \zeta_+},$$

*where $c$ is independent of the particular sequence $h_n$, of $K$, $\zeta$ and $\zeta_+$.*

The second lemma is a corollary of the Cauchy integral formula:

**Cauchy's Estimate.** *Let $h(z)$ be an analytic function, defined on a domain $U \subset \mathbf{C}$. Let $z_0 \in U$ and $V = \{z : |z - z_0| < \rho\} \subset U$, and let $\sup_V |h(z)| \leqq K$. Then*

$$\left| h^{(n)}(z_0) \right| \leqq \frac{n!}{\rho^n} K.$$

### *4.2. The conjugacy*

Here Proposition 1, used in Section 3.3.6 is proved. The proof will follow from two lemmas. The first provides the solution of the (19). The second deals with (21) and (22).

**4.2.1. The classical homological equation.** For the following lemma, see for instance [1].

**Lemma 6.** *Let $\omega \in \mathcal{A}'_c + \rho$. Let $\zeta > \zeta_+ > 0$. Let $h$ be analytic on $\left(S^1 + \zeta\right) \times (\mathcal{A}'_c + \rho)$ and assume that there $|h| \leqq \delta$ holds. Assume also that $\rho \leqq \frac{1}{2}\gamma d^{-\tau}$.*

*If it is required that $\int_{S^1} U(\xi, \sigma) \mathrm{d}\xi = 0$, then the equation*

$$U(\xi + \omega, \sigma) - U(\xi, \sigma) = {}_d h(\xi, \sigma) + w(\sigma) \tag{29}$$

*has a unique solution $(U, w)$. On $\left(S^1 + \zeta_+\right) \times (\mathcal{A}'_c + \rho)$ we have*

$$|U|_{\zeta_+} \leqq c' \frac{\delta}{(\zeta - \zeta_+)^{\tau+2}},$$

*where $c' > 1$ depends only on $\tau$. On $\mathcal{A}'_c + \rho$ we have*

$$|w| \leqq \delta.$$

**Proof.** The lemma is proved using Fourier analysis. If

$$h(\xi, \sigma) = \sum_{n \in \mathbf{Z}} h_n(\sigma) \mathrm{e}^{in\xi}$$

then the trigonometric polynomial:

$$U(\xi, \sigma) = U_0(\sigma) + \sum_{0<n \leqq d} \frac{h_n(\sigma)}{\mathrm{e}^{in\omega} - 1} \mathrm{e}^{in\xi}, \tag{30}$$

solves (29), provided the denominators do not vanish. The condition:

$$\int_{S^1} U = 0$$

implies $U_0(\sigma) = 0$ for all $\sigma$, and, by integrating both sides of (29) over $S^1$,

$$w(\sigma) = -h_0(\sigma). \tag{31}$$

Thus (29) has at most one solution.

A necessary and sufficient condition for the existence of a solution is that the small divisors $\mathrm{e}^{in\omega} - 1$ do not vanish (recall that $U$ is a trigoniometric polynomial). Since $\omega \in \mathcal{A}'_c + \rho$, we have that $\left|\omega - 2\pi \frac{p}{q}\right| \geqq \gamma q^{-\tau} - \rho$ for all $p \in \mathbf{Z}$, $q \in \mathbf{Z}\backslash\{0\}$. The denominators can now be estimated as follows:

$$\begin{aligned} \left|\mathrm{e}^{in\omega} - 1\right| &= \left|\mathrm{e}^{i(n\omega - 2\pi p)} - 1\right| \\ &\geqq \frac{2}{\pi} \left|\omega - 2\pi \frac{p}{n}\right| |n| \\ &\geqq \frac{2}{\pi} \left(\gamma |n|^{-\tau} - \rho\right) |n|. \end{aligned}$$

The integer $p$ is chosen such that $\omega - 2\pi p \in (-\pi, \pi]$ – this determines $p \in \mathbf{Z}$ uniquely – and then it is used that $|\mathrm{e}^{ix} - 1| \geqq 2|x|/\pi$ if $x \in [-\pi, \pi]$. By the Paley-Wiener lemma, we have the following estimate on the $h_n(\sigma)$:

$$|h_n(\sigma)| \leqq 2\pi |h|_\zeta \mathrm{e}^{-|n|\zeta}.$$

This implies the following estimate on $U$:

$$\begin{aligned} |U(\xi, \sigma)|_{\zeta_+} &\leqq \sum_{0<n \leqq d} \frac{2\pi |h|_\zeta \mathrm{e}^{-|n|\zeta}}{\frac{2}{\pi}\left(\gamma |n|^{-\tau} - \rho\right) |n|} \mathrm{e}^{|n|\zeta_+} \\ &\leqq\cdot \ |h|_\zeta \sum_{0<n \leqq d} \frac{|n|^{\tau+1} \mathrm{e}^{-|n|(\zeta - \zeta_+)}}{1 - \rho \frac{d^\tau}{\gamma}} \\ &\leqq\cdot \ \frac{|h|_\zeta}{(\zeta - \zeta_+)^{\tau+2}} \int_0^\infty |x|^{\tau+1} \mathrm{e}^{-x} \mathrm{d}x \leqq\cdot \ \frac{|h|_\zeta}{(\zeta - \zeta_+)^{\tau+2}}. \end{aligned}$$

In the third inequality we used the fact that $\rho \leqq \frac{1}{2}\gamma d^{-\tau}$, and we made the substitution $|n|(\zeta - \zeta_+) \sim x$.

The estimate of $|w|$ is simply

$$|w(\sigma)| = |h_0(\sigma)| \leqq 2\pi |h|_\zeta \leqq\cdot\ |h|_\zeta .$$

This completes the proof of the lemma □

**4.2.2. Twisting behaviour normal to the torus.**

**Lemma 7.** *Let $\omega \in \mathcal{A}+\rho$ and $\beta \in \{x \in \mathbf{R}^2 : |x| = 1\}+\alpha$. Let $0 < \zeta_+ < \zeta \leqq 1/k$ (where $k \in \mathbf{Z}$ is the "twisting constant").*

*Let $\Sigma = (\mathcal{A} + \rho) \times (\{z : |z| = 1\} + \alpha)$. Let $h(\xi, \sigma) = (h^1, h^2)$ be real analytic and assume that $h$ can be analytically extended to $\left(S^1 + \zeta\right) \times \Sigma$; denote the complex analytical extension also by $h$. Assume that for the extension*

$$|h|_\zeta \leqq K$$

*holds. Assume also that $\alpha \leqq k\zeta/4$ and $\rho \leqq d^{-2}$. Let $v$ denote the vector valued function $(v_1(\xi, \sigma), v_2(\xi, \sigma))$. Then the equation*

$$v(\xi + \omega, \sigma) - E(\xi, \beta)v(\xi, \sigma) = {}_d h(\xi, \sigma) + \sum_{\ell=0}^{k-1} \begin{pmatrix} \cos \ell\xi & -\sin \ell\xi \\ \sin \ell\xi & \cos \ell\xi \end{pmatrix} \begin{pmatrix} w_{\ell,1} \\ w_{\ell,2} \end{pmatrix} \tag{32}$$

*has a real analytic solution $(v_1, v_2, w_{0,1}, \ldots, w_{k-1,2})$. On $\left(S^1 + \zeta_+\right) \times \Sigma$ we have*

$$|v|_{\zeta_+} \leqq\cdot\ \frac{K}{(\zeta - \zeta_+)\zeta}$$

*and*

$$|w| \leqq\cdot\ \frac{K}{\zeta}.$$

Note that we will not encounter small divisor problems. The truncation of the function $h$ on the right-hand side of (32) is necessary in order to have uniform estimates with respect to $\omega$.

**Proof.** Since $h^1(\xi, \sigma)$ and $h^2(\xi, \sigma)$ are real analytic, by the Paley-Wiener lemma their Fourier coefficients satisfy the following inequalities:

$$\left|h_n^s(\sigma)\right| \leqq 2\pi |h|_\zeta \mathrm{e}^{-|n|\zeta} \ \text{ for } s = 1, 2. \tag{33}$$

Equation (32) is re-written, using functions $H$, $V$ and $W$, given by

$$\begin{aligned} H(\xi, \sigma) &= {}_d h^1(\xi, \sigma) + i\,{}_d h^2(\xi, \sigma), \\ V(\xi, \sigma) &= v^1(\xi, \sigma) + i v^2(\xi, \sigma), \\ W_n(\sigma) &= w_{n1}(\sigma) + i w_{n2}(\sigma), \end{aligned}$$

and using the parameter $b = \beta_1 + i\beta_2$. Note that $H$, $V$ and $W$ do not necessarily take real values on real vectors. By adding the first component of (32) to $i$ times the second component, and by re-writing the equation using the new functions $H$, $V$ and $W$, we arrive at the following equation:

$$V(\xi+\omega,\sigma) - b\mathrm{e}^{ik\xi} V(\xi,\sigma) = H(\xi,\sigma) + \sum_{\ell=0}^{k-1} W_\ell(\sigma)\mathrm{e}^{i\ell\xi}. \tag{34}$$

Put $V(\xi,\sigma) = \sum V_n(\sigma)\mathrm{e}^{in\xi}$ etc. Then from (34) the following infinite system of coupled linear equations is obtained:

$$V_n(\sigma)\mathrm{e}^{in\omega} - bV_{n-k}(\sigma) = H_n(\sigma) + \sum_{\ell=0}^{k-1} W_\ell(\sigma)\delta_{n\ell}. \tag{35}$$

Here $\delta_{mn}$ is the Kronecker delta: $\delta_{mn} = 0$ if $m \neq n$ and $\delta_{mn} = 1$ if $m = n$. As an aside, note that for $k = 0$ this would be a "small divisor" equation.

There are $k$ "independent" series of coupled equations

$$(\ldots, V_{-k}, V_0, V_k, V_{2k}, \ldots), \quad (\ldots, V_{-k+1}, V_1, V_{k+1}, V_{2k+1}, \ldots) \quad \text{etc.}$$

Each of these series can be solved independently of the others. To consider one series of equations, the integer $\ell$ is fixed, $n = km + \ell$ is taken, and $m$ is used as a new variable. The coefficients $\tilde{V}_m^\ell$, $\tilde{H}_m^\ell$ are introduced by

$$\begin{aligned} \tilde{V}_m^\ell(\sigma) &= V_{km+\ell}(\sigma) \\ \tilde{H}_m^\ell(\sigma) &= H_{km+\ell}(\sigma). \end{aligned} \tag{36}$$

Then (35) takes the form (dropping the argument $\sigma$ and the superposed $^\ell$)

$$\begin{aligned} \tilde{V}_m\mathrm{e}^{i(km+\ell)\omega} - b\tilde{V}_{m-1}, &= \tilde{H}_m & m &\neq 0, \\ \tilde{V}_0\mathrm{e}^{i\ell\omega} - b\tilde{V}_{-1} &= \tilde{H}_0 + W_\ell, & m &= 0. \end{aligned} \tag{37}$$

This series of equations can be interpreted as a difference equation for the Fourier coefficients $\tilde{V}_m$, in the sense that one "initial condition" $\tilde{V}_{\bar{m}}$ for some $\bar{m}$ determines all other $\tilde{V}_m$.

In order that the Fourier series $\sum V_n\mathrm{e}^{in\xi}$ converges, one should have the absolute values of the coefficients $|V_n|$, $|V_{-n}|$ tending to 0 as $n \to \infty$; consequently it is necessary that $|\tilde{V}_m|$, $|\tilde{V}_{-m}| \to 0$ as $m \to \infty$.

Call $\tilde{d} = \text{Entier}(\frac{d-\ell}{k}) + 1$. For $m > \tilde{d}$, we see that $|\tilde{V}_{m+1}| = |b||\tilde{V}_m|$, since $\tilde{H}_m = 0$ for $m > \tilde{d}$. Since we should have that $|V_n| \to 0$ as $n \to \infty$, uniformly in $|b|$, it follows that $\tilde{V}_m = 0$ for all $m \geqq \tilde{d}$. Starting from $V_{d+1}$, $\tilde{V}_m$ can be determined for $0 \leqq m \leqq \tilde{d}$ from the difference equations (37). Likewise, $\tilde{V}_{-m} = 0$ for all $m > \tilde{d}$, and $\tilde{V}_{-m}$ can be computed recursively for $1 \leqq m \leqq \tilde{d}$.

At this level, we can understand the necessity of the "parameter shift" $W_\ell$. For generic choices of $\{\tilde{H}_m\}_{m\in\mathbf{Z}}$, we have that:

$$\tilde{V}_0\mathrm{e}^{i\ell\omega} - b\tilde{V}_{-1} \neq \tilde{H}_0.$$

In other words, the equation

$$V(\xi+\omega,\sigma) - b\mathrm{e}^{ik\xi} V(\xi,\sigma) = H(\xi,\sigma)$$

generically has no solution $V(\xi, \sigma)$ in the class of continuous functions. To remove that *obstruction*, the parameter $W_\ell$ is introduced by

$$W_\ell = \tilde{V}_0 e^{i\ell\omega} - b\tilde{V}_{-1} - \tilde{H}_0.$$

This way a solution to the difference equation (37) is constructed, unique in the class of solutions were $|\tilde{V}_m| \to 0$ as $|m| \to \infty$.

For precise estimates, some additional information is needed. So to determine the solutions of (37) explicitely, put $\tilde{V}^u_{\tilde{d}} = \tilde{V}_{\tilde{d}} = 0$ and determine $\tilde{V}^u_{\tilde{d}-1}$, $\tilde{V}^u_{\tilde{d}-2}$ etc. recursively. Likewise, put $\tilde{V}^l_{-\tilde{d}} = \tilde{V}_{-\tilde{d}} = 0$, and determine $\tilde{V}^l_{-(\tilde{d}-1)}$, $\tilde{V}^l_{-(\tilde{d}-2)}$ etc. This yields (some tedious, but straightforward algebra is omitted here; however it is feasible to verify the solution by substitution in (37))

$$\begin{aligned}
\tilde{V}^u_m &= -b^m \exp\left(-i\frac{1}{2}m\omega(2\ell + k(m+1))\right) \\
&\qquad \sum_{r=m+1}^{\tilde{d}} b^{-r}\tilde{H}_r(\sigma)\exp\left(i\frac{1}{2}(r-1)\omega(2\ell + kr)\right) \\
\tilde{V}^l_{-m} &= b^{-m}\exp\left(i\frac{1}{2}(m-1)\omega(2\ell - km)\right) \\
&\qquad \sum_{r=m}^{\tilde{d}} b^r \tilde{H}_{-r}(\sigma)\exp\left(i\frac{1}{2}r\omega((r+1)k - 2\ell)\right)\ ;
\end{aligned}$$

in both cases $m \geqq 0$.

The Fourier coefficients $\tilde{V}^u_m$ and $\tilde{V}^l_{-m}$ have to be estimated. First estimate $|\tilde{H}_r|$ using (33):

$$\begin{aligned}
|\tilde{H}_r| = |H_{kr+\ell}| &\leqq \cdot \ \ |h|_\zeta e^{-|kr+\ell|\zeta} \\
&\leqq \cdot \ \ |h|_\zeta e^{-k|r|\zeta}.
\end{aligned}$$

Also sums of the form $\sum_r b^r e^{-r\zeta}$ have to be estimated (for all $\zeta$ between 0 and $1/k$). Since by the assumptions of this lemma $|b| < 1+\alpha$, $\alpha \leqq \frac{k\zeta}{4}$ and $k\zeta \leqq 1$, it follows that

$$\begin{aligned}
|be^{-k\zeta}| &\leqq \left(1 + \frac{k\zeta}{4}\right)\left(1 - k\zeta + \frac{(k\zeta)^2}{2}\right) \\
&= 1 - \frac{3}{4}k\zeta + \frac{(k\zeta)^2}{4} + \frac{(k\zeta)^3}{8} \\
&\leqq 1 - \frac{k\zeta}{4}
\end{aligned}$$

and

$$
\begin{aligned}
|b^{-1}\mathrm{e}^{-k\zeta}| &\leqq \frac{1-k\zeta+\frac{(k\zeta)^2}{2}}{1-\frac{k\zeta}{4}} \\
&\leqq \frac{1-\frac{k\zeta}{4}}{1-\frac{k\zeta}{4}} + \frac{-\frac{3}{4}k\zeta+\frac{1}{2}k\zeta}{\frac{3}{4}} \\
&\leqq 1-\frac{k\zeta}{3}.
\end{aligned}
$$

From these inequalities, it follows that $\sum_{r=0}^{\infty}(b\mathrm{e}^{-k\zeta})^r$ and $\sum_{r=0}^{\infty}(b^{-1}\mathrm{e}^{-k\zeta})^r$ converge; for instance

$$
\begin{aligned}
\sum_{r=0}^{\infty}\left(b\mathrm{e}^{-k\zeta}\right)^r &\leqq \sum_{r=0}^{\infty}\left(1-\frac{k\zeta}{8}\right)^r \\
&\leqq \frac{1}{1-\left(1-\frac{k\zeta}{8}\right)} \\
&\leqq\cdot\ \frac{1}{\zeta};
\end{aligned}
$$

likewise for $\sum_{r=0}^{\infty}(b^{-1}\mathrm{e}^{-k\zeta})^r$. Using this, the fact that $\omega \in \mathcal{A}+\rho$ and the relation $\rho < d^{-2}$, $|\tilde{V}_m^u|$ is estimated:

$$
\begin{aligned}
|\tilde{V}_m^u| &\leqq b^m \exp\left(\frac{m\rho}{2}(2\ell+km+k)\right) \\
&\qquad \sum_{r=m+1}^{d} b^{-r}\left|\tilde{H}_r(\sigma)\right| \exp\left(\frac{(r-1)\rho}{2}(2\ell+kr)\right) \\
&\leqq\cdot\ b^m \mathrm{e}^{k\rho m^2}\mathrm{e}^{k\rho d^2}\sum_{r=m+1}^{d} b^{-r}|h|_\zeta \mathrm{e}^{-(kr+\ell)\zeta} \\
&\leqq\cdot\ |h|_\zeta\, \mathrm{e}^{2k\rho d^2}\mathrm{e}^{-(km+\ell)}\sum_{r=0}^{\infty}\left(b\mathrm{e}^{k\zeta}\right)^{-r} \\
&\leqq\cdot\ \frac{|h|_\zeta}{\zeta}\mathrm{e}^{-(km+\ell)}.
\end{aligned}
$$

In the same way $|\tilde{V}_m^l|$ is estimated. We conclude that

$$
|V_n(\sigma)| \leqq\cdot\ \frac{|h|_\zeta}{\zeta}\mathrm{e}^{-|n|\zeta} \text{ for all } n \in \mathbf{Z} \text{ and} \tag{38}
$$

$$
|W_\ell| \leqq\cdot\ \frac{|h|_\zeta}{\zeta}. \tag{39}
$$

This proves already that

$$
|w| \leqq\cdot\ \frac{K}{\zeta}.
$$

Now the results of the $k$ uncoupled infinite systems of equations are pieced together, and a solution $V$ is found where

$$V(\xi,\sigma) = \sum_{k=-d}^{d} V_n \mathrm{e}^{in\xi}.$$

Since $V(\xi,\sigma)$ is a trigonometric polynomial in $\xi$, and since the $H_n(\sigma)$ are analytic in $\sigma$, the function $V(\xi,\sigma)$ is analytic in its variables. By the second part of the Paley-Wiener lemma, on $S^1 \times \zeta_+$ the following estimate for $V(\xi,\sigma)$ holds:

$$|V(\xi,\sigma)|_{\zeta_+} \leqq \cdot \ \frac{|h|_\zeta}{\zeta(\zeta-\zeta_+)}. \tag{40}$$

If we restrict ourselves to the real line, by real analyticity real and imaginary parts can be taken and one arrives at

$$v_1(\xi,\sigma) = \mathrm{Re}\ V(\xi,\sigma) = \sum_{\substack{k=-d\\ k\neq 0}}^{d} \mathrm{Re}\ V_n(\sigma)\cos n\xi - \mathrm{Im}\ V_n(\sigma)\sin n\xi$$

$$v_2(\xi,\sigma) = \mathrm{Im}\ V(\xi,\sigma) = \sum_{\substack{k=-d\\ k\neq 0}}^{d} \mathrm{Re}\ V_n(\sigma)\sin n\xi + \mathrm{Im}\ V_n(\sigma)\cos n\xi.$$

The functions $v_1(\xi,\sigma)$ and $v_2(\xi,\sigma)$ are real analytic functions on real vectors; thus they can be analytically extended to complex vectors in a unique way.

Likewise

$$w_{\ell,1}(\sigma) = \mathrm{Re}\ W_\ell(\sigma),$$
$$w_{\ell,2}(\sigma) = \mathrm{Im}\ W_\ell(\sigma).$$

The estimates of the $v_i$ and $w_{\ell,i}$ where $i=1,2$, follow from (40) and (39) respectively. □

Note that the neighbourhoods $S^1+r_j$ of $S^1$ contain the neighbourhood $S^1+r_0/2$, so that, when applying the lemma, $\zeta$ is always bounded from below by $r_0/2$ and therefore dissappears in generic constants.

**4.2.3. Application to the equations.**

**Proof of the Proposition 1.** Turning to the solution ofs (19), (21) and (22), on $\mathcal{D}$

$$|f| + \frac{|g|}{s} \leqq \delta$$

is assumed. Starting with (19),

$$u(\xi+\omega,\sigma) - u(\xi,\sigma) = w_1(\sigma) + {}_d f(\xi,0,\sigma),$$

first $|u|$ will be estimated on $\mathcal{D}_{\frac{1}{4}}$. This domain is chosen because $u$ appears on the right-hand side of (21) and (22). Application of Lemma 7 requires some extra “space” in the $\xi$-direction.

Since $\rho$ and $d$ are such that $\rho = \frac{1}{2}\gamma d^{-\tau}$ (see Section 3.3.5), Lemma 6 can be applied. We get that

$$|u|_{\frac{1}{4}} \leqq c' \frac{\delta}{(r - r_{\frac{1}{4}})^{\tau+2}} \leqq \cdot\ 4^{j\tau}\delta, \tag{41}$$

and

$$|w_1(\sigma)| \leqq \delta. \tag{42}$$

Note that the first inequality of (41), together with inequality (27), implies

$$r_{\frac{1}{4}} + |u|_{\frac{1}{4}} \leqq r_{\frac{1}{4}} + \frac{r - r_{\frac{1}{4}}}{4^{j\tau}} \leqq r \tag{43}$$

Turn to (21):

$$\begin{aligned} &v_0(\xi + \omega, \sigma) - E(\xi, \beta)v_0(\xi, \sigma) = \\ &\qquad M(\xi, w_3) + {}_d g(\xi, 0, \sigma) + {}_d \left\{M(\xi + u, \mu) - M(\xi, \mu)\right\}. \end{aligned}$$

This equation is of the form treated in Lemma 7, with

$$h(\xi, \sigma) = g(\xi, 0, \sigma) + M\Big(\xi + u(\xi, \sigma), \mu\Big) - M(\xi, \mu).$$

By the mean value theorem,

$$|M(\xi + u, \mu) - M(\xi, \mu)|_{\frac{1}{4}} \leqq \sup_{S^1+r} \left|\frac{\partial M}{\partial \xi}\right| |u|_{\frac{1}{4}} \leqq \cdot\ |\mu| 4^{j\tau}\delta$$

since $r_{\frac{1}{4}} + |u|_{\frac{1}{4}} \leqq r$ by (43). Also

$$|h| \leqq \cdot\ 4^{j\tau} s\delta,$$

since $|\mu| \leqq s$. Applying Lemma 7, it follows for $v_0$ and $w_3$ that

$$|v_0|_{\frac{1}{2}} \leqq \cdot\ \frac{4^{j\tau} s\delta}{(r_{\frac{1}{4}} - r_{\frac{1}{2}})} \leqq \cdot\ 8^{j\tau} s\delta \tag{44}$$

and

$$|w_3|_{\frac{1}{2}} \leqq \cdot\ 4^{j\tau} s\delta. \tag{45}$$

Finally turn to (22):

$$\begin{aligned} &v_1(\xi + \omega, \sigma)E(\xi, \beta) - E(\xi, \beta)v_1(\xi, \sigma) \\ &\qquad = E(\xi, w_2) + L(\xi, w_4) + {}_d g_\eta(\xi, 0, \sigma) \\ &\qquad\quad + {}_d \Big\{E(\xi + u, \beta) - E(\xi, \beta) + L(\xi + u, \lambda) - L(\xi, \lambda)\Big\}. \end{aligned}$$

The equation is multiplied from the right with $E(\xi, \beta)^{-1}$ to obtain

$$\begin{aligned} & v_1(\xi+\omega,\sigma) - E(\xi,\beta)v_1(\xi,\sigma)E(\xi,\beta)^{-1} \\ & \quad = E(\xi, w_2)E(\xi,\beta)^{-1} + L(\xi, w_4)E(\xi,\beta)^{-1} \\ & \qquad +_d g_\eta(\xi, 0, \sigma)E(\xi,\beta)^{-1} +_d \Big\{ E(\xi+u,\beta) - E(\xi,\beta) \\ & \qquad + L(\xi+u,\lambda) - L(\xi,\lambda)\Big\} E(\xi,\beta)^{-1}. \end{aligned} \tag{46}$$

Here some work has to be done. Set

$$V(\xi,\sigma) = \begin{pmatrix} V_{11}(\xi,\sigma) & V_{12}(\xi,\sigma) \\ V_{21}(\xi,\sigma) & V_{22}(\xi,\sigma) \end{pmatrix} \equiv v_1(\xi,\sigma)$$

and

$$\begin{aligned} H(\xi,\sigma) = {} & g_\eta(\xi,0,\sigma)E(\xi,\beta)^{-1} \\ & + \Big\{ E(\xi+u,\beta) - E(\xi,\beta) + L(\xi+u,\lambda) - L(\xi,\lambda)\Big\} E(\xi,\beta)^{-1}. \end{aligned}$$

Note that $V(\xi,\sigma)$ and $H(\xi,\sigma)$ are $2\times 2$-matrices. Recall that any $2\times 2$ matrix $T = (t_{ij})$ can be split into a conformal orientation preserving part (for short “conformal”), and a conformal orientation reversing part (for short “anti-conformal”). Referring to Section 2.2,

$$\begin{aligned} T &= \begin{pmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} t_{11}+t_{22} & t_{12}-t_{21} \\ -t_{12}+t_{21} & t_{11}+t_{22} \end{pmatrix} + \frac{1}{2}\begin{pmatrix} t_{11}-t_{22} & t_{12}+t_{21} \\ t_{12}+t_{21} & -t_{11}+t_{22} \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} t_{11}+t_{22} & t_{12}-t_{21} \\ -t_{12}+t_{21} & t_{11}+t_{22} \end{pmatrix} + \frac{1}{2}\begin{pmatrix} t_{11}-t_{22} & -t_{12}-t_{21} \\ t_{12}+t_{21} & t_{11}-t_{22} \end{pmatrix} C \\ &= T^c + T^{ac}C. \end{aligned}$$

Here $C = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Note that both $T^c$ and $T^{ac}$ are conformal.

Conformal orientation preserving maps commute. Moreover, we have that

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} C = C \begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

The definitions of Section 2.2 imply that the linear mapping $E(x,b)$ is conformal, while $L(x,\ell)$ is anti-conformal. Now, let

$$V = V^c + V^{ac}C = \begin{pmatrix} A_1 & -A_2 \\ A_2 & A_1 \end{pmatrix} + \begin{pmatrix} B_1 & -B_2 \\ B_2 & B_1 \end{pmatrix} C.$$

Likewise, let

$$H = H^c + H^{ac}C = \begin{pmatrix} F_1 & -F_2 \\ F_2 & F_1 \end{pmatrix} + \begin{pmatrix} G_1 & -G_2 \\ G_2 & G_1 \end{pmatrix} C.$$

Define $A = (A_1, A_2)$, etc. Let $\psi$ be defined by $\cos\psi = \beta_1/\|\beta\|$, $\sin\psi = \beta_2/\|\beta\|$. Moreover, let $\hat{\beta} = (\cos 2\psi, \sin 2\psi)$.

Then (46) splits up into a conformal and an anti-conformal part; using the above definitions, we get

$$A(\xi+\omega) - A(\xi) = \frac{2}{\|\beta\|}\begin{pmatrix} \cos\psi & \sin\psi \\ -\sin\psi & \cos\psi \end{pmatrix}\begin{pmatrix} w_{2;1} \\ w_{2;2} \end{pmatrix} + F \tag{47}$$

$$\begin{aligned} &B(\xi+\omega) - E(2\xi,\hat{\beta})B(\xi) = \\ &\qquad \frac{2}{\|\beta\|}\sum_{n=0}^{2k-1}\begin{pmatrix} \cos n\xi & -\sin n\xi \\ \sin n\xi & \cos n\xi \end{pmatrix}\begin{pmatrix} \cos\psi & -\sin\psi \\ \sin\psi & \cos\psi \end{pmatrix}\begin{pmatrix} w_{4;n,1} \\ w_{4;n,2} \end{pmatrix} + G. \end{aligned} \tag{48}$$

The equation for $A$ is of the same form as in Lemma 6, and the equation for $B$ is of the same form as in Lemma 7, if one takes there $2k$ instead of $k$. Estimate $F$ and $G$:

$$\begin{aligned} |F|_{\frac14}, |G|_{\frac14} \leqq\cdot\ |H|_{\frac14} &\leqq\cdot\ \left|g_\eta(\xi,0,\sigma)\right|_{\frac14} + |E(\xi+u,\beta) - E(\xi,\beta)|_{\frac14} \\ &\qquad + |L(\xi+u,\lambda) - L(\xi,\lambda)|_{\frac14} \\ &\leqq\cdot\ \frac{|g|}{s} + \sup_{S^1+r}\left(\left|\frac{\partial E}{\partial\xi}\right| + \left|\frac{\partial L}{\partial\xi}\right|\right)|u|_{\frac14} \\ &\leqq\cdot\ \delta + 4^{j\tau}\delta \leqq\cdot\ 4^{j\tau}\delta. \end{aligned}$$

Equations (47) and (48) can be solved now using the two lemmas; then we transform back to $(V, w_2, w_4)$ and from there to $(v_1, w_2, w_4)$ to obtain solutions

$$|v_1|_{\frac12} \leqq\cdot\ 8^{j\tau}\delta, \tag{49}$$

$$|w_2| + |w_4| \leqq\cdot\ 4^{j\tau}\delta. \tag{50}$$

Combining the estimates (41), (42), (44), (45), (49) and (50) yields the statement of the proposition. □

### *4.3. The remainders*

Here we prove Proposition 3 from Section 3.3.6.

**Proof of Proposition 3.** The proof of this proposition will be split up into a series of lemmas, where terms of the same kind will be treated simultaneously. In these lemmas, implicit use will be made of the definitions of the domains in Section 3.3.5, as well as of the estimates given by Proposition 3.

**4.3.1. Shift of variables.** Here the terms $t_1$ and $t_4$ are going to be treated. Recall that

$$t_1 = f(\xi + u, \eta + v, \sigma + W) - f(\xi, \eta, \sigma),$$
$$t_4 = g(\xi + u, \eta + v, \sigma + W) - g(\xi, \eta, \sigma).$$

Since the mean value theorem will be applied, the derivatives of $f$ and $g$ have to be estimated.

**Lemma 8.** *Let* (28) *hold. Then we have on* $\mathcal{D}_{\frac{1}{2}}$ *that*

$$\left|\frac{\partial f}{\partial \xi}\right|_{\frac{1}{2}} + \frac{1}{s}\left|\frac{\partial g}{\partial \xi}\right|_{\frac{1}{2}} \leqq\cdot\ 2^j\delta, \quad \left|\frac{\partial f}{\partial \eta}\right|_{\frac{1}{2}} + \frac{1}{s}\left|\frac{\partial g}{\partial \eta}\right|_{\frac{1}{2}} \leqq\cdot\ \frac{\delta}{s},$$
$$\left|\frac{\partial f}{\partial \omega}\right|_{\frac{1}{2}} + \frac{1}{s}\left|\frac{\partial g}{\partial \omega}\right|_{\frac{1}{2}} \leqq\cdot\ \frac{\delta}{\rho}, \quad \left|\frac{\partial f}{\partial \beta}\right|_{\frac{1}{2}} + \frac{1}{s}\left|\frac{\partial g}{\partial \beta}\right|_{\frac{1}{2}} \leqq\cdot\ 2^j\delta,$$
$$\left|\frac{\partial f}{\partial \mu}\right|_{\frac{1}{2}} + \frac{1}{s}\left|\frac{\partial g}{\partial \mu}\right|_{\frac{1}{2}} \leqq\cdot\ \frac{\delta}{s}, \quad \left|\frac{\partial f}{\partial \lambda}\right|_{\frac{1}{2}} + \frac{1}{s}\left|\frac{\partial g}{\partial \lambda}\right|_{\frac{1}{2}} \leqq\cdot\ \frac{\delta}{q}.$$

**Proof.** This is a repeated application of Cauchy's estimate, taking into account the precise forms of the domains $\mathcal{D}$ and $\mathcal{D}_{\frac{1}{2}}$. □

**Lemma 9.** *Let (28) hold. Then we have:*

$$|t_1|_+ + \frac{|t_4|_+}{s} \leqq\cdot\ 8^{j\tau}\frac{\delta^2}{q}.$$

**Proof.** Estimating $t_1$ and $t_4$ is a straightforward application of the mean value theorem. The magnitude of

$$|f(\xi + u, \eta + v, \omega + w_1, \beta + w_2, \mu + w_3, \lambda + w_4) - f(\xi, \eta, \omega, \beta, \mu, \lambda)|_+$$

has to be estimated, as well as the same expression for $g$ instead of $f$. This yields

$$\begin{aligned}
&|f(\xi + u, \eta + v, \omega + w_1, \beta + w_2, \mu + w_3, \lambda + w_4) - f(\xi, \eta, \omega, \beta, \mu, \lambda)|_+\\
&\quad\leqq\cdot\ \left|\frac{\partial f}{\partial \xi}\right|_{\frac{1}{2}} |u|_+ + \left|\frac{\partial f}{\partial \eta}\right|_{\frac{1}{2}} |v|_+\\
&\qquad + \left|\frac{\partial f}{\partial \omega}\right|_{\frac{1}{2}} |w_1|_+ + \left|\frac{\partial f}{\partial \beta}\right|_{\frac{1}{2}} |w_2|_+ + \left|\frac{\partial f}{\partial \mu}\right|_{\frac{1}{2}} |w_3|_+ + \left|\frac{\partial f}{\partial \lambda}\right|_{\frac{1}{2}} |w_4|_+\\
&\quad\leqq\cdot\ 2^j\delta \cdot 4^{j\tau}\delta + \frac{\delta}{s}\cdot 4^{j\tau}s\delta + \frac{\delta}{\rho}\cdot\delta + 2^j\delta\cdot 2^{j\tau}\delta + \frac{\delta}{s}\cdot s\delta + \frac{\delta}{q}\cdot 2^{j\tau}\delta\\
&\quad\leqq\cdot\ 8^{j\tau}\frac{\delta^2}{q}
\end{aligned}$$

making heavy use of the previous lemma and the fact that $q < \rho$ and that $\tau > 2$. Likewise it follows that

$$|g(\xi+u, \eta+v, \omega+w_1, \beta+w_2, \mu+w_3, \lambda+w_4) - g(\xi, \eta, \omega, \beta, \mu, \lambda)|_+ \leqq\cdot\ 8^{j\tau}s\frac{\delta^2}{q}$$

which proves the lemma. □

**4.3.2. Linearization.** Turning to the terms $t_2$ and $t_5$, recall that

$$t_2 = f(\xi, \eta, \sigma) - f(\xi, 0, \sigma),$$
$$t_5 = g(\xi, \eta, \sigma) - g(\xi, 0, \sigma) - g_\eta(\xi, 0, \sigma)\eta.$$

The estimates for $|t_2|$ is as above in Lemma 9, and $|t_5|$ is treated with the corresponding estimate, where $g_{\eta\eta}$ is estimated with Cauchy. The result is stated in the following lemma.

**Lemma 10.** *Let* (28) *hold. Then we have*

$$|t_2|_+ \leqq\cdot \frac{s_+}{s}\delta, \qquad \frac{|t_5|_+}{s} \leqq\cdot \frac{s_+^2}{s^2}\delta.$$

**4.3.3. Truncation.** Here the terms $t_3$, $t_6$, $t_{12}$, $t_{13}$ and $t_{14}$ are estimated. These are the errors made by truncating the Fourier series. Recall that

$$\begin{aligned}
t_3 &= f(\xi, 0, \sigma) - {}_d f(\xi, 0, \sigma), \\
t_6 &= g(\xi, 0, \sigma) + g_\eta(\xi, 0, \sigma) - {}_d \left\{ g(\xi, 0, \sigma) + g_\eta(\xi, 0, \sigma) \right\}, \\
t_{12} &= M(\xi + u, \mu) - M(\xi, \mu) - {}_d \left\{ M(\xi + u, \mu) - M(\xi, \mu) \right\}, \\
t_{13} &= \left\{ E(\xi + u, \beta) - E(\xi, \beta) \right\} \eta - {}_d \left\{ E(\xi + u, \beta) - E(\xi, \beta) \right\} \eta, \\
t_{14} &= \left\{ L(\xi + u, \lambda) - L(\xi, \lambda) \right\} \eta - {}_d \left\{ L(\xi + u, \lambda) - L(\xi, \lambda) \right\} \eta.
\end{aligned}$$

First we prove the following lemma. The estimates follow as a simple corollary.

**Lemma 11.** *Let* $0 < \zeta_+ < \zeta < 1$. *Let* $h$ *be real analytic on* $S^1 + \zeta$. *Then we have on* $S^1 + \zeta_+$ *that*

$$|h - {}_d h|_+ \leqq\cdot |h| \frac{\mathrm{e}^{-d(\zeta - \zeta_+)}}{\zeta - \zeta_+}.$$

**Proof.** Let $\sum_{n=-\infty}^{\infty} h_n \mathrm{e}^{in\xi}$ be the Fourier series of $h$. By the Paley-Wiener lemma it follows that

$$|h_n| \leqq 2\pi |h| \mathrm{e}^{-|n|\zeta}.$$

Estimate $|h - {}_d h|_+$:

$$\begin{aligned}
|h - {}_d h|_+ &= \left| \sum_{|n| > d} h_n \mathrm{e}^{in\xi} \right|_+ \leqq\cdot \sum_{|n| > d} |h_n| \mathrm{e}^{|n|\zeta_+} \\
&\leqq\cdot |h| \sum_{|n| > d} \mathrm{e}^{-|n|(\zeta - \zeta_+)} \leqq\cdot |h| \frac{\mathrm{e}^{-d(\zeta - \zeta_+)}}{\zeta - \zeta_+},
\end{aligned}$$

which was to be shown. □

**Corollary 12.** *The terms* $t_3$, $t_6$, $t_{13}$, $t_{14}$ *and* $t_{15}$ *are such that*

$$|t_3|_+ + \frac{|t_6|_+}{s} + \frac{|t_{13}|_+}{s} + \frac{|t_{14}|_+}{s_+} + \frac{|t_{15}|_+}{s_+ q} \leqq\cdot 2^j \mathrm{e}^{-d(r - r_+)}\delta. \tag{51}$$

**4.3.4. Other terms.** Here the rest of the $|t_i|$'s: $i = 7, \dots, 12$ will be estimated. Recall that

$$\begin{aligned} t_7 &= M(\xi + u, w_3) - M(\xi, w_3), \\ t_8 &= \{E(\xi + u, w_2) - E(\xi, w_2)\}\,\eta, \\ t_9 &= \{E(\xi + u, \beta + w_2) - E(\xi, \beta)\}\,v, \\ t_{10} &= \{L(\xi + u, w_4) - L(\xi, w_4)\}\,\eta, \\ t_{11} &= \{L(\xi + u, \lambda + w_4) - L(\xi, \lambda)\}\,v, \\ t_{12} &= L(\xi + u, \lambda)v. \end{aligned}$$

These terms are treated in the three subsequent lemmas, each dealing with a group of terms of the same kind:

**Lemma 13.** *The terms $t_7$, $t_8$ and $t_{10}$ are such that*

$$|t_7|_+ + |t_8|_+ + |t_{10}|_+ \leqq\cdot\ 4^{j\tau} s\delta^2.$$

**Proof.** Take $|t_7|_+$:

$$\begin{aligned} |M(\xi + u, w_3) - M(\xi, w_3)|_+ &\leqq\cdot\ |u|_+ |w_3|_+ \\ &\leqq\cdot\ 4^{j\tau} s\delta^2. \end{aligned}$$

The terms $|t_8|_+$ and $|t_{10}|_+$ can be treated together:

$$\begin{aligned} &\left|\Big\{E(\xi + u, w_2) - E(\xi, w_2)\Big\}\eta\right|_+ + \Big|\{L(\xi + u, w_4) - L(\xi, w_4)\}\,\eta\Big|_+ \\ &\qquad \leqq\cdot\ |u|_+|w_2|_+|\eta|_+ + |u|_+|w_4|_+|\eta|_+ \\ &\qquad \leqq\cdot\ 2^{j\tau} s_+ \delta^2. \end{aligned}$$

Since $s_+ < s$, combination with the estimate of $|t_7|_+$ yields the result. □

**Lemma 14.** *The terms $t_9$ and $t_{11}$ are such that*

$$|t_9|_+ + |t_{11}|_+ \leqq\cdot\ 8^{j\tau} s\delta^2.$$

**Proof.** The terms $|t_9|_+$ and $|t_{11}|_+$ can be estimated in together:

$$\begin{aligned} &\Big|\{E(\xi + u, \beta + w_2) - E(\xi, \beta)\}\,v\Big|_+ \\ &\qquad + \Big|\{L(\xi + u, \lambda + w_4) - L(\xi, \lambda)\}\,v\Big|_+ \\ \leqq &\qquad \Big|E(\xi + u, \beta + w_2) - E(\xi + u, \beta)\Big|_+ |v|_+ \\ &\qquad + \Big|E(\xi + u, \beta) - E(\xi, \beta)\Big|_+ |v|_+ \\ &\qquad + \Big|L(\xi + u, \lambda + w_4) - L(\xi + u, \lambda)\Big|_+ |v|_+ \\ &\qquad + \Big|L(\xi + u, \lambda) - L(\xi, \lambda)\Big|_+ |v|_+ \\ \leqq\cdot &\quad (2|u|_+ + |w_2|_+ + |w_4|_+)\,|v|_+ \\ \leqq\cdot &\quad 8^{j\tau} s\delta^2. \qquad \square \end{aligned}$$

**Lemma 15.** *The term $t_{12}$ is such that*

$$|t_{12}|_+ \leqq \cdot \; q_+ 4^{j\tau} s\delta.$$

**Proof.** We have

$$|L(\xi+u,\lambda)v|_+ \leqq \cdot \; |\lambda|_+ |v|_+ \leqq \cdot \; q_+ 4^{j\tau} s\delta. \qquad \square$$

### *4.4. Finishing the induction*

Here we prove Proposition 4 from Section 3.3.6.

**Proof of Proposition 4.** The expression $|f^+|_+ + |g^+|_+/s_+$ is estimated using Proposition 3 in Section 3.3.6, that is,

$$\begin{aligned} |f^+|_+ + \frac{|g^+|_+}{s_+} &\leqq 2\left(\sum_{i=1}^{3} |t_i|_+ + \frac{1}{s_+}\sum_{i=4}^{17} |t_i|_+\right) \\ &\leqq 2c_2\Bigg(8^{j\tau}\frac{\delta^2}{q} + \frac{s_+}{s}\delta + 2^j \mathrm{e}^{-d(r-r_+)}\delta + 8^{j\tau}\frac{s}{s_+}\frac{\delta^2}{q} \\ &\qquad + \frac{s_+}{s}\delta + 2^j\frac{s}{s_+}\mathrm{e}^{-d(r-r_+)}\delta + 8^{j\tau}\frac{s}{s_+}\delta^2 \\ &\qquad + 4^{j\tau}\frac{s}{s_+}\delta q_+ + 8^{j\tau}s\delta|f^+|_+\Bigg). \end{aligned}$$

Using inequality (27) and the fact that the $s_j$ are decreasing, we estimate the factor $8^{j\tau}s\delta$ in front of $|f^+|_+$ on the right-hand side. Since $\delta_j$ converges faster to 0 than any geometrical sequence, we have that $8^{j\tau}\delta_j$ is bounded. As the sequence $\{s_j\}_{j=0}^{\infty}$ is decreasing, it follows that

$$2c_2 \cdot 8^{j\tau}s\delta \leqq Ks_0,$$

for some $K$ depending only on $r_0$ and $c_2$. Choose $s_0 < (2K)^{-1}$. Then

$$2c_2 8^{j\tau}s\delta|f^+|_+ \leqq \frac{1}{2}|f^+|_+,$$

and

$$\begin{aligned} &|f^+|_+ + \frac{|g^+|_+}{s} \\ &\quad \leqq \cdot \; 8^{j\tau}\delta\left(\frac{\delta}{q}\left(1+\frac{s}{s_+}\right) + \frac{s_+}{s} + \mathrm{e}^{-d(r-r_+)}\left(1+\frac{s}{s_+}\right) + \frac{s}{s_+}\big(\delta+q_+\big)\right). \end{aligned}$$

To find sequences $\{\delta_j\}_j$, $\{s_j\}_j$ and $\{q_j\}_j$ such that the induction "works", the following *Ansatz* is tried (see Section 3.3.5):

$$\delta_+ = \delta^{\kappa}, \quad \frac{s_+}{s} = \delta^{\zeta_1} \quad \text{and} \quad q = \delta^{\zeta_2},$$

where $\zeta_1, \zeta_2 > 0$ are undetermined constants and have nothing to do with any $\zeta$'s defined previously; the constant $\kappa$ is required to be larger than 1 in order to have fast convergence of the $\delta_j$.

Note that because $\delta_0 < 1$, the $s_j$ are strictly decreasing, and thus $s/s_+ > 1$. Also the $q_j$ tend to zero, and actually $q_j < 1$ for all $j$.

The last expression can be simplified to:

$$|f^+|_+ + \frac{|g^+|_+}{s_+} \leqq c_3 8^{j\tau} \left( \delta \frac{s}{s_+} \frac{\delta}{q} + \delta \frac{s}{s_+} q_+ + \delta \frac{s_+}{s} + \delta \mathrm{e}^{-d(r-r_+)} \frac{s}{s_+} \right), \tag{52}$$

where $c_3$ only depends on $\tau, \gamma, k$ and $r_0$. The right-hand side has to be made smaller than $\delta^\kappa$. This is done by choosing $\kappa$, $\zeta_1$ and $\zeta_2$ in such a way that the first three terms,

$$\delta \frac{s}{s_+} \frac{\delta}{q} + \delta \frac{s}{s_+} q_+ + \delta \frac{s_+}{s},$$

of the right-hand side of (52) are smaller than $\delta^{\kappa+\varepsilon}$, with $\varepsilon > 0$. (The extra factor $\delta^\varepsilon$ is needed to deal with the factor $c_3 8^{j\tau}$). Then it will be shown that by the choice made for the sequence $\{d_j\}$ in Section 3.3.5, the fourth term

$$\delta \mathrm{e}^{-d(r-r_+)} \frac{s}{s_+}$$

is also smaller than $\delta^{\kappa+\varepsilon}$ as well.

The *Ansatz* yields the conditions:

$$\begin{aligned} \delta\delta^{-\zeta_1} \frac{\delta}{\delta^{\zeta_2}} &\leqq \delta^{\kappa+\varepsilon}, \\ \delta\delta^{-\zeta_1} \delta^{\kappa\zeta_2} &\leqq \delta^{\kappa+\varepsilon}, \\ \delta\delta^{\zeta_1} &\leqq \delta^{\kappa+\varepsilon}, \end{aligned}$$

or simply (note that $\delta < 1$):

$$\begin{aligned} 2 - \zeta_1 - \zeta_2 &\geqq \kappa + \varepsilon, \\ 1 - \zeta_1 + \kappa\zeta_2 &\geqq \kappa + \varepsilon, \\ 1 + \zeta_1 &\geqq \kappa + \varepsilon. \end{aligned}$$

These inequalities can be fulfilled if the choices $\zeta_1 = \frac{1}{4}, \zeta_2 = \frac{1}{2}, \kappa = \frac{6}{5}$ and $\varepsilon = \frac{1}{20}$ are made.

It remaines to compute

$$\delta \mathrm{e}^{-d(r-r_+)} \frac{s}{s_+}.$$

Recall from (3.3.5) that $r - r_+ = r_0 2^{-j-2}$ and

$$d = (2\kappa)^j \frac{2}{r_0} \log \delta_0^{-1}.$$

Moreover,

$$\delta = \delta_0^{\kappa^j} = \delta_0^{\left(\frac{6}{5}\right)^j}. \tag{53}$$

Substitution yields

$$\delta \mathrm{e}^{-d(r-r_+)} \frac{s}{s_+} = \delta_0^{\frac{3}{4}\left(\frac{6}{5}\right)^j} \mathrm{e}^{-\frac{1}{2}\left(\frac{6}{5}\right)^j \log \delta_0^{-1}}$$
$$= \delta^{\frac{5}{4}} = \delta^{\kappa+\varepsilon}.$$

So (52) is now of the form

$$|f^+|_+ + \frac{|g^+|_+}{s_+} \leqq c_3 8^{j\tau} \delta^{\varepsilon} \cdot \delta^{\kappa}.$$

Consider $\log\left(c_3 8^{j\tau}\delta^{\varepsilon}\right)$:

$$\log\left(c_3 8^{j\tau}\delta^{\varepsilon}\right) = \log c_3 + j\tau \log 8 + \varepsilon\kappa^j \log \delta_0.$$

This is less than 0 for all $j \in \{0, 1, 2, \ldots\}$ if $\delta_0$ is chosen small enough. But then

$$c_3 8^{j\tau}\delta^{\varepsilon} \leqq 1$$

for all $j$, and

$$|f^+|_+ + \frac{|g^+|_+}{s_+} \leqq \delta^{\kappa} = \delta_+.$$

That finishes the proof of the proposition. □

### *4.5. Applying the inverse approximation*

In this subsection Proposition 5 of Section 3.3.7 is proved.

**Proof of Proposition 5.** Note that all the diffeomorphism $\Phi_j$ and $\mathcal{U}_j$ are close to the identity. Moreover, the $\Phi_j$ (as well as the $\mathcal{U}_j$) are all of the form $\Phi(x, y, p) = \Phi_0(x, p) + (0, \Phi_1(x, p)y, 0)$. Thus

$$\Phi \circ \mathcal{U} = \Phi_0 \circ \pi_{(x,p)} \circ \mathcal{U} + (0, \left(\Phi_1 \circ \pi_{(x,p)} \circ \mathcal{U}\right) \pi_y \circ \mathcal{U}, 0),$$

where $\pi_y(x, y, p) = y$, $\pi_{(x,p)}(x, y, p) = (x, p)$ (below $\pi_x(x, y, p) = x$ will be used). In order to estimate $|\Phi_j|$, first expand $\Phi_j$:

$$\begin{aligned}
\Phi_j &= \left(x + \bar{u}^j(x, p), y + \bar{v}_0^j(x, p) + \bar{v}_1^j(x, p)y, p + \bar{W}^j(p)\right) \\
&= \mathcal{U}_0 \circ \mathcal{U}_1 \circ \cdots \circ \mathcal{U}_{j-1} \\
&= \Big(\pi_x(\mathcal{U}_1 \circ \cdots \circ \mathcal{U}_{j-1}) + u^0 \circ \pi_{(x,p)} \circ \mathcal{U}_1 \circ \cdots \circ \mathcal{U}_{j-1}, \\
&\qquad \pi_y(\mathcal{U}_1 \circ \cdots \circ \mathcal{U}_{j-1}) + v_0^0 \circ \pi_{(x,p)} \circ \mathcal{U}_1 \circ \cdots \circ \mathcal{U}_{j-1} \\
&\qquad + \left(v_1^0 \circ \pi_{(x,p)} \circ \mathcal{U}_1 \circ \cdots \circ \mathcal{U}_{j-1}\right) \pi_y(\mathcal{U}_1 \circ \cdots \circ \mathcal{U}_{j-1}), \\
&\qquad \pi_p(\mathcal{U}_1 \circ \cdots \circ \mathcal{U}_{j-1}) + W^0 \circ \pi_{(x,p)} \circ \mathcal{U}_1 \circ \cdots \circ \mathcal{U}_{j-1}\Big)
\end{aligned}$$

$$= \Bigg( x + \sum_{n=0}^{j-1} u^n \circ \pi_{(x,p)} \circ \mathcal{U}_{n+1} \circ \cdots \circ \mathcal{U}_{j-1},$$

$$y + \sum_{m=0}^{j-1} \left( \prod_{n=0}^{m-1} v_1^n \circ \pi_{(x,p)} \circ \mathcal{U}_{n+1} \circ \cdots \circ \mathcal{U}_{j-1} \right) \cdot v_0^m \circ \pi_{(x,p)} \circ \mathcal{U}_{m+1} \circ \cdots \circ \mathcal{U}_{j-1}$$

$$+ \left( \prod_{n=0}^{j-1} v_1^n \circ \pi_{(x,p)} \circ \mathcal{U}_{n+1} \circ \cdots \circ \mathcal{U}_{j-1} \right) y,$$

$$p + \sum_{n=0}^{j-1} W^n \circ \pi_{(x,p)} \circ \mathcal{U}_{n+1} \circ \cdots \circ \mathcal{U}_{j-1} \Bigg).$$

Lemma 16 will prove useful for the estimates that follow.

**Lemma 16.** *Fix constants $c, M, m > 0$. Then there is an $N > 0$, such that for $n > N$, the following inequality holds:*

$$c\, \mathrm{e}^{Mn^m} \delta_n \leqq \left(\frac{1}{2}\right)^n.$$

**Proof.** Since $\delta_n = \delta_0^{\left(\frac{6}{5}\right)^n}$ and $\log \delta_0 < 0$, for $n$ sufficiently large it follows that

$$\log c + Mn^m + \left(\frac{6}{5}\right)^n \log \delta_0 < -n \log 2. \qquad \square$$

The functions $\bar{u}^j$, $\bar{v}^j$ and $\bar{W}^j$ have to be estimated on $\mathcal{D}_j$. Using the result of Proposition 1, we obtain

$$|\bar{u}^j(x, p)|_{\mathcal{D}_j} \leqq \sum_{n=0}^{j-1} |u^n|_{\mathcal{D}_{n+1}} \leqq \sum_{n=0}^{\infty} c_1 4^{n\tau} \delta_n.$$

This sum converges because of the lemma. Likewise it is established that the $|\bar{v}_n^j|_{\mathcal{D}_j}$ ($n = 0, 1$) and $|\bar{w}_n^j|_{\mathcal{D}_j}$ ($n = 1, 2, 3, 4$) are uniformly bounded.

To apply the inverse approximation lemma (see Appendix A), the difference

$$\left|\Phi_{j+1} - \Phi_j\right|_{\mathcal{D}_{j+1}}$$

has to be estimated. Since $\Phi_{j+1} = \Phi_j \circ \mathcal{U}_j$, the difference reads (see Sections (3.2.2) and (3.3.2)):

$$\Phi_{j+1} - \Phi_j = \left( u^j + \bar{u}^j \circ \mathcal{U}_j - \bar{u}^j, v^j + \bar{v}^j \circ \mathcal{U}_j - \bar{v}^j, W^j + \bar{W}^j \circ \mathcal{U}_j - \bar{W}^j \right).$$

This gives

$$
\begin{aligned}
|\Phi_+ - \Phi|_+ \leqq & \left( \left(1 + \left|\frac{\partial \bar{u}}{\partial x}\right|_{\frac{3}{4}}\right) |u|_+ + \left|\frac{\partial \bar{u}}{\partial p}\right|_{\frac{3}{4}} |W|_+, \right.\\
& \left|\frac{\partial \bar{v}}{\partial x}\right|_{\frac{3}{4}} |u|_+ + (1 + |\bar{v}_1|_{\frac{3}{4}})|v|_+ + \left|\frac{\partial \bar{v}}{\partial p}\right|_{\frac{3}{4}} |W|_+,\\
& \left. \left(1 + \left|\frac{\partial \bar{W}}{\partial p}\right|_{\frac{3}{4}}\right) |W|_+ \right)\\
\leqq & M\delta^{\frac{1}{2}}.
\end{aligned}
$$

In the first estimate, the mean value theorem is used together with the fact that

$$\mathcal{U}(\mathcal{D}_+) \subset \mathcal{D}_{\frac{3}{4}}$$

(see Corollary 2). In the second estimate the derivatives of $\Phi_j$ have been estimated using Cauchy's estimate; the $|u|_+$, etc. again have been estimated using Proposition 1. A factor $\delta^{\frac{1}{2}}$ has been sacrificed to control the factors $4^{j\tau}$ arising from Cauchy's estimate on geometrically shrinking domains. The constant $M$ is independent of $j$.

The Inverse Approximation Lemma can be applied to obtain a unique function $\Phi_\infty$ on $\mathcal{D}_\infty$, which is $C^\infty$. Moreover, for all $m > 0$, in the $C^m$ norm $\|.\|_m$, we have that

$$\|\Phi_\infty - \Phi_j\|_m \to 0.$$

From this it follows that for a fixed value of the parameter $p$, the convergence of the diffeomorphisms $\Phi_j(.,.,p)$ to $\Phi_\infty(.,.,p)$ is uniform. Convergence in the $x$-direction is on an open piece of $(\mathbf{C}/2\pi\mathbf{Z})$; in the $y$-direction the $\Phi_j$ are affine, so that for fixed parameter $p$, the map $\Phi_\infty(.,.,p)$ is even real analytic.

Now the conjugacy property of $\Phi_\infty$ is evident from its construction. This finishes the proof of the proposition. □

*Acknowledgements.* We wish to thank FLORIS TAKENS for reading the manuscript thoroughly many times, as well as for valuable discussions and remarks. Also, we wish to thank HANS DUISTERMAAT, HANS DE JONG, HEINZ HANSSMANN, ROBERT MACKAY, ANATOLI NEISHTADT, JÜRGEN PÖSCHEL and MIKHAEL SEVRYUK for fruitful comments and discussions.

## Appendix A. The inverse approximation lemma

In this appendix, the inverse approximation lemma is stated, in a version taylored down from the formulation in [2] (see also [14], [17]).

**Inverse Approximation Lemma.** *Let $\beta > 0$, $\beta \notin \mathbf{N}$, and $m = \sup\{n \in \mathbf{N} : n < \beta\}$; $\rho_j = ab^j$ with $a > 0$, $0 < b < 1$; moreover, let $\Omega \subset \mathbf{R}$ be a closed set, and*

$$W_j = \Omega + \rho_j = \bigcup_{x \in \Omega} \{z \in \mathbf{C} \mid |z - x| < \rho_j\}.$$

*Let* $\left\{\Psi^j\right\}_{j=0}^{\infty}$ *be such that* $\Psi^j$ *is real analytic on* $W^j$*, and for* $j \geqq 0$

$$|\Psi^{j+1} - \Psi^j|_{W_j} \leqq M\rho_j^{\beta}$$

*for some constant* $M$.

*Then there exists a unique function* $\Psi^{\infty}$*, defined on* $\Omega$*, which is extendable to a* $C^m$ *function on* $W_0$*, and which is such that*

$$\|\Psi^{\infty}\|_m \leqq Mc_{\beta},$$

*where* $c_{\beta}$ *only depends on* $a$ *and* $b$.

*Moreover,*

$$\|\Psi^{\infty} - \Psi^j\|_m \to 0 \text{ as } j \to \infty.$$

## References


1. V. I. Arnol'd, *Geometric methods in the theory of ordinary differential equations*, Springer, 1983.
2. B. L. J. Braaksma & H. W. Broer, On a quasi-periodic Hopf Bifurcation, *Analyse non linéaire* **4** (1987), 115–168.
3. H. W. Broer, G. B. Huitema & M. Sevryuk, *Quasi-periodic motions in families of dynamical systems: order amidst chaos*, LNM **1645** (1996).
4. H. W. Broer, G. B. Huitema, F. Takens & B. L. J. Braaksma, Unfoldings and bifurcations of quasi-periodic tori, *Memoirs of the AMS* **83** (1990), 1–175.
5. H. W. Broer & F. Takens, Formally symmetric normal forms and genericity, *Dynamics Reported*, **2** (1989), 39–59.
6. H. W. Broer & F. Takens, Mixed spectra and rotational symmetry, *Arch. Rat. Mech. Anal.* **124** (1993), 13–42.
7. H. W. Broer, F. Takens & F. O. O. Wagener, Integrable and non-integrable deformations of the skew Hopf bifurcations, *Reg. and Chaot. Dyn.* **4** (1999), 17–43
8. A. Chenciner & G. Iooss, Bifurcations des tores invariantes, *Arch. Rational Mech. Anal.* **69** (1979), 109–198.
9. A. Chenciner & G. Iooss, Persistance et bifurcations de tores invariants, *Arch. Rat. Mech. Anal.* **71** (1979), 301–306.
10. J. Guckenheimer & P. Holmes, *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*, Springer, 1983.
11. H. Hanssmann, The quasi-periodic centre-saddle bifurcation, *J. Differential Equations* **142** (1998), 305–370.
12. M. W. Hirsch, C. Pugh & M. Shub, *Invariant manifolds*, Lecture Notes in Math. **583** (1977).
13. J. Pöschel, Integrability of Hamiltonian systems on Cantor sets, *Comm. Pure and Appl. Math.*, **35** (1982), 653–696.
14. E. Stein, *Singular integrals and differentiability-properties of functions*, Princeton, 1970.
15. F. O. O. Wagener, *On the skew Hopf bifurcation*, Thesis, University of Groningen 1998.
16. H. Whitney, Analytic extensions of differentiable functions defined in closed sets, *Trans. Amer. Math. Soc.* **36** (1934), 63–89.

17. E. ZEHNDER, Generalized implicit function theorems with applications to some small divisor problems I, *Comm. Pure Appl. Math.* **28** (1975), 91–140.

Department of Mathematics
University of Groningen
Groningen
The Netherlands

and

Department of Economics
University of Amsterdam
Amsterdam
The Netherlands

(*Accepted September 3, 1999*)